Travel Time Perception Errors: Causes and Consequences

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Carlos Carrion

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor Of Philosophy

David M. Levinson

July 2013

Acknowledgement

Dedico hac dissertatione ad Deum.

In nomine Patris, et Filii, et Spiritus Sancti. Amen.

Gloria Patri, et Filio, et Spiritui Sancto. Sicut erat in principio, et nunc, et semper, et in saecula saeculorum. Amen.

I want to express my most sincere appreciation and respect to Dr. David Levinson. His support, creativity, and encouragement has allowed me to finish this dissertation. In addition, He has encouraged me to pursue my interests in mathematical economics, and econometrics. I am thankful for his guidance through the course of my graduate work.

I would also like to thank professors (Drs. Harder, Liu, Cao, Davis, Roe, Coggins, Apland, Homans, and Glewwe) for conversations and insight during my time here at the University of Minnesota. I have learned quite a bit about urban planning, mathematical optimization, econometrics, microeconomics, psychology, and many other topics thanks to them. I also thank my fellow graduate students (Dr. Zhu, Dr. Tilahun, Dr. Parthasarathi, Arthur, Indrajit, Xuan, Saif, Avital, and Henry), and other researchers (Michael) for their friendship, and support that further contributed to my work.

Last, but not least, I thank my loved ones for their continuous and unconditional support in all my endeavors regardless of their nature.

ABSTRACT

This research investigates the causes, and consequences behind travel time perception. Travel times are *experienced*. Thus, travelers estimate the travel time through their own *perception*. This is the underlying reason behind the mismatch between travel times as reported by a traveler (*subjective travel time distribution*) and travel times as measured from a device (e.g. loop detector or GPS navigation device; *objective travel time distribution*) in collected data. It is reasonable that the relationship between subjective travel times and objective travel times may be expressed mathematically as: $T_s = T_o + \xi$. T_s is a random variable associated with the probability density given by the *subjective travel time distribution*. T_o is a random variable associated with the probability density given by the *objective travel time distribution*. The variable ξ is the random *perception error* also associated with its own probability density. Thus, it is clear that travelers may overestimate or underestimate the measured travel times, and this is likely to influence their decisions unless $\mathrm{E}(\xi) = 0$, and $\mathrm{Var}(\xi) \approx 0$. In other words, travelers are "optimizing" (i.e. executing decisions) according to their own divergent views of the *objective travel time distribution*.

This dissertation contributes novel results to the following areas of transportation research: travel time perception; valuation of travel time; and route choice modeling. This study presents a systematic identification of factors that lead to perception errors of travel time. In addition, the factors are related to similar factors on time perception research in psychology. These factors are included in econometric models to study their influence on travel time perception, and also identify which of these factors lead to overestimation or underestimation of travel times. These econometric models are estimated on data collected from commuters recruited from a previous research study in the Minneapolis-St. Paul region (Carrion and Levinson, 2012a, Zhu, 2010). The data (surveys, and Global Positioning System [GPS] points) consists of work trips (from home to work, and from work to home) of subjects. For these work trips, the subjects' self-reported travel times, and the subjects' travel times measured by GPS devices were collected. Furthermore, this dissertation

provides the first empirical results that highlight the influence of perception errors in the valuation of travel time, and in the dynamic behavior of travelers' route choices. Last but not least important, this dissertation presents the most comprehensive literature review of the value of travel time reliability written to date.

Contents

List of Tables

List of Figures

List of Variables

Chapter 1

β	parameters to be estimated
σ^2	variance of random perception error
ξ	random perception error
$E(\cdot)$	expected value operator
$f(\cdot)$	functional form of the covariates
$Var(\cdot)$	variance operator
T_o	random objective travel time
T_s	random subjective travel time
$\mathbf{x}$	vector of covariates

Chapter 2

β, γ	parameters to be estimated
λ_k	degree of risk aversion
μ_T	centrality measure of travel time
σ_T	dispersion measure of travel time
$E(\cdot)$	expected value operator
$Var(\cdot)$	variance operator
C	travel cost
DL	delay indicator

P_{AB}	route choice set between origin pair AB
P_L	probability of arriving late
PAT	preferred arrival time
RR	reliability ratio
SDE	schedule delay early
SDL	schedule delay late
T	random travel time
T_p	random travel time of path p
VOR	value of travel time reliability
VOT	value of travel time savings
$VSDE$	value of scheduling delay early
$VSDL$	value of scheduling delay late

Chapter 4

β, γ	parameters to be estimated
$E(\cdot)$	expected value operator
$Prob[\cdot]$	cumulative probability function
$f(\cdot)$	functional form of the covariates
$Var(\cdot)$	variance operator
$\mathcal{N}$	set of subjects
n	index for a generic subject
$\mathcal{J}_n$	set of observations for subject n
j	index for a generic observation
τ_{jn}	ratio of reported travel time to measured travel time. See equation 4.1
δ_{jn}	indicator for overestimated vs underestimated travel time. See equation 4.3
T_{mjn}	measured travel time from GPS data. See equation 4.1
T_{rjn}	reported travel time from survey data. See equation 4.1
$\mathbf{x}$	vector of continuous covariates described in section 4.2.3
$\mathbf{z}$	vector of categorical covariates described in section 4.2.3

ϵ_{jn}	$\sim$ i.i.d. $N(0, \sigma_\epsilon^2)$ for all j and n
α_n	$\sim$ i.i.d. $N(0, \sigma^2)$ for all n
l_P	length (km) of trip along a path P
l_{la}	length (km) of trip on limited access roads along a path P
l_{sa}	length (km) of trip on signalized arterials along a path P
l_d	euclidean distance (km) of trip along a path P
l_e	network distance (km; trip length) of trip along a path P
C_P	circuity of path P
Y_P	discontinuity of the trip along a path P
A_P	proportion of signalized arterials of path P
H_P	proportion of limited access roads of path P

Chapter 5

β, γ	parameters to be estimated
ξ	random perception error
$Cov(\cdot)$	covariance operator
$E(\cdot)$	expected value operator
$f(\cdot)$	functional form of the covariates
$Var(\cdot)$	variance operator
T_o	random objective travel time
T_s	random subjective travel time
$\mathcal{N}$	set of subjects
n	index for a generic subject
$\mathcal{C}_n$	set of choices for subject n
j	index for a generic alternative
U_{nj}	random utility of subject n for alternative j
V_{nj}	systematic utility of subject n for alternative j
ϵ_{nj}	$\sim$ i.i.d. Gumbel distribution(0,1) for all j and n
$\mathbf{x}$	vector of covariates

RR	reliability ratio
VOR	value of travel time reliability
VOT	value of travel time savings

Chapter 6

β, γ	parameters to be estimated
$Prob[\cdot]$	cumulative probability function
$f(\cdot)$	functional form of the covariates
n	index for a generic subject
T_n	single-spell duration of subject n
$F_n(t)$	cumulative probability function of T_n
$f_n(t)$	probability density function of T_n
$S_n(t)$	survivor function of T_n
$h_n(t)$	hazard function of T_n
$H_n(t)$	cumulative hazard function of T_n
$\mathbf{x}$	vector of covariates
$\mathbf{y}$	vector of time-invariant covariates
$\mathbf{w}(\mathbf{t})$	vector of time-dependent covariates
$\phi(x, \beta)$	relative hazard function of T_n
$h_0(t)$	baseline hazard function of T_n

Chapter 1

Introduction

Travel time is an indispensable characteristic of any transportation system. It is an important pillar that shapes the decisions of travelers (i.e. the demand side) in the transportation market. Typically, travelers must choose within reason (or at least within their own "reason"): where to locate for residence and in cases for work; whether a trip is necessary or it can be postponed; a destination and purpose (e.g. grocery shopping); and when to schedule a trip; which mode to select, and dependent on the mode, which route to choose. Perhaps even other choice dimensions may be present but to a lesser extent (e.g. whether to group activities or not; how fast to travel). These choices denote the complexity of the travelers choice problem where travel time is an influential element.

There are two aspects of travel time that affect traveler's choices: perception, and predictability. Perception refers to traveler's very own interpretation of the time elapsed during their trips. This interpretation is based on many factors linked to their memory (i.e. previous experiences), and senses (i.e. interpretation of travel time passage). Predictability refers to the uncertainty of travel time due to several factors such as: heterogeneity of drivers and their vehicles; traffic regulations, and traffic management systems; traffic incidents (e.g. traffic signal failure, vehicular crashes); weather patterns; and others. Predictability is also related to perception as travelers previous experiences with uncertainty allow them to estimate the duration of their trips, and new experiences with uncertainty will become part of the accumulated experiences of travelers.

My research aims to study the causes and consequences of perception errors of travel time. It is inspired by psychological research (Allan, 1979, Madalina, 2011) that indicates travelers may overestimate or underestimate the duration of their activities depending on several factors (i.e. causes of perception errors) including but not limited to: the effort of performing the activity; the value they assign to the activity; the relative priority (i.e. working vs. shopping) of the activity; previous experiences; and others. In addition, this research extends Parthasarathi (2011), and Parthasarathi et al. (2012) where difference between measured (i.e. obtained from mechanical instruments) travel time and perceived (i.e. obtained from travelers statements) travel time is linked with the built environment. The consequences of perception errors are studied in two cases: valuation of travel time used in economic analyses; and dynamic behavior of commuters used in route choice modeling. The underlying principle of both cases is that travelers are optimizing according to their own divergent view of the actual travel time distribution (i.e. measured from devices). Consequently, travelers will differ in their optimal solutions depending on the degree of distortion (i.e. error) of their perception of the actual distribution of travel time. This is a significant contribution to the field as currently no study has attempted to connect (empirically) the influences of the perception of travel time to the valuation of travel time, and to the dynamic behavior of commuters.

Lastly, this research exploits a unique data source of collected GPS devices and surveys (Zhu, 2010, Carrion and Levinson, 2012a). There are no other datasets that contain commute information at this level of detail for a span of 8-10 weeks for each subject. Figure 1.1 shows the number of observations per link in the Minneapolis-St. Paul road network. This illustrates the extensive coverage of the GPS data set. The main characteristics of the data are that actual route information as in spatial location of subjects in the network is known, home and work location of subjects are known, and also subjects filled several question regarding their travel experience, and their time restrictions.

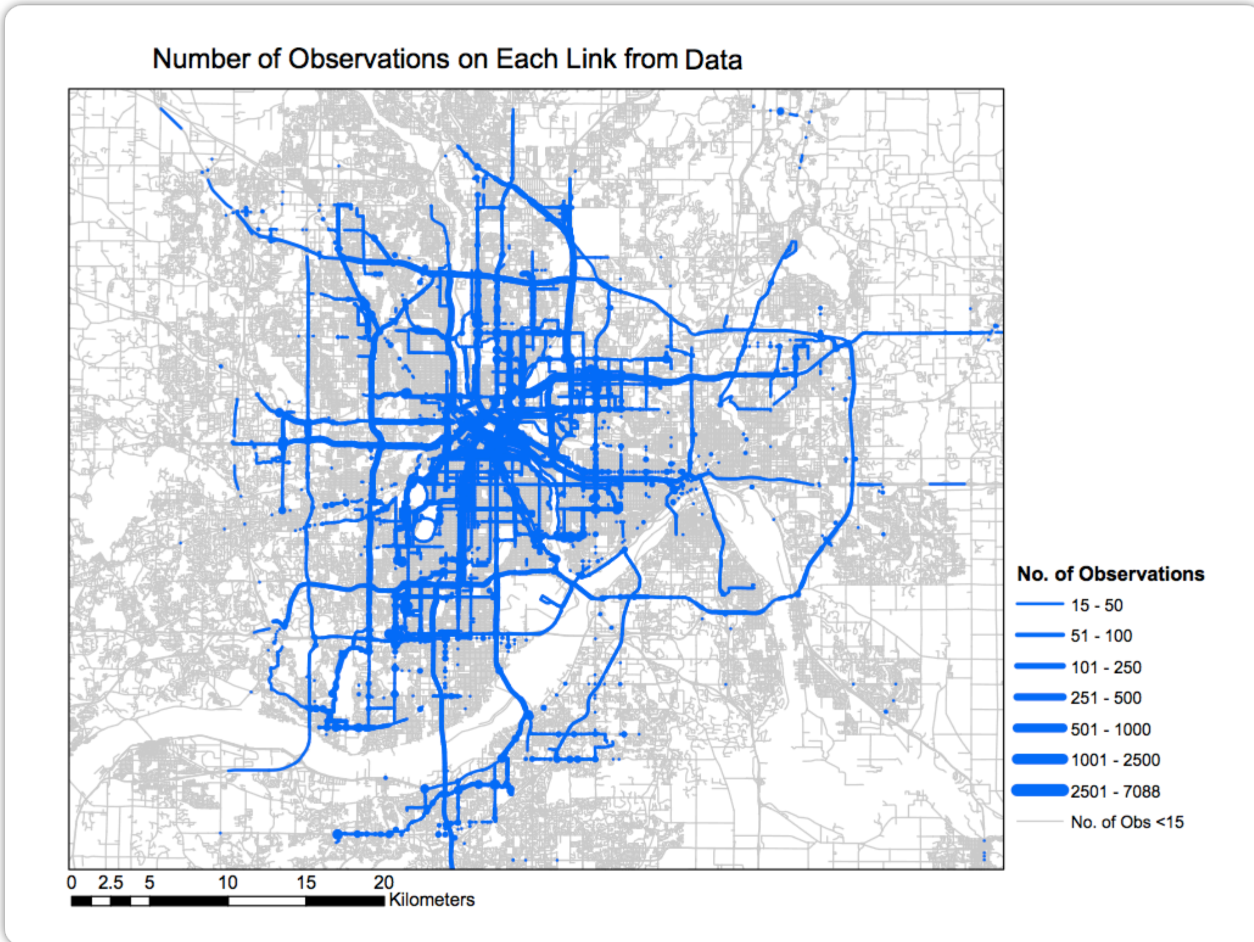


Figure 1.1: Number of observations per link for GPS data set (Source: Zhu (2010)).

1.1 The role of perception errors of travel time in auto commute trips

In transportation research, there are studies (Levinson et al., 2004, 2006, Peer et al., 2010, Peer, 2013) indicating that travelers overestimate or underestimate the travel times of their trips. Thus, it should be well known to researchers that travelers are responding to travel times according to their own perceptions. Nevertheless, this knowledge has not been reflected extensively in the research literature, where travel time is such an important input in many modeling efforts. It is true that there are theoretical efforts in areas of transportation research such as route choice modeling. However, the empirical evidence is lacking, and it is mostly concentrated on travel time perception of travelers without significant interaction with other areas such as valuation of travel time.

Travel times are *experienced*. Travelers estimate the travel time of their trips through their own cognitive mechanism of *perception*. The *perception errors* in the mechanism are the underlying reason behind the mismatch between travel times as reported by a traveler (i.e. *subjective travel time distribution*) and travel times as measured from a device (e.g. loop detector or GPS navigation device; i.e. *objective travel time distribution*) in collected data. The term *perception errors* refers to the distortions in the mechanism of the travelers not allowing them to estimate accurately the actual experienced travel time. A useful explanation of the *role of perception errors* is formulated as follows: $T_s = T_o + \xi$. T_s is a random variable associated with the probability density given by the *subjective travel time distribution*. T_o is a random variable associated with the probability density given by the *objective travel time distribution*. The random variable ξ represents *perception errors*. This variable is also associated with its own probability density. Thus, it is clear that travelers may overestimate or underestimate the measured travel times, and this is likely to influence their decisions unless $\mathrm{E}(\xi) = 0$, and $\mathrm{Var}(\xi) \approx 0$. However, it is likely that $\mathrm{E}(\xi) \neq 0$, and most likely $\mathrm{E}(\xi) = f(\mathbf{x}; \beta)$ (see chapter 4). Thus, the *perception errors* (ξ) are not purely by chance, but are also influenced by different factors summarized in the vector $\mathbf{x}$ with its respective parameters β. In addition, the $\mathrm{Var}(\xi)$ is likely to not be approximately zero. It is plausible that $\mathrm{Var}(\xi)$ is homoskedastic ($\mathrm{Var}(\xi) = \sigma^2$; see chapter 4). In some cases, the

Var(ξ) could be heteroskedastic (Var(ξ) $= f(\mathbf{x}; \beta)$), but for the purposes of this illustration it is sufficient just to acknowledge that it is likely that it is not approximately zero.

Psychologists have studied time perception far longer than transportation researchers, and they have classified the perception of time into three main categories: subjective time passage (i.e. perception of the speed that time passes); estimation of time duration; and simultaneity and succession of time. The estimation of time duration is the most frequently studied category by psychologists, and thus it is better understood. It is also the dimension of time perception that will be focus of this study, and it has been the focus of most studies investigating perception of travel time in the transportation literature. Main factors identified in the duration of time are: temporal relevance, temporal uncertainty, affective elements, arousal, task complexity, temporal expectancies, absorption and attentional deployment. *Temporal relevance* refers to the significance of time for performing a task in an optimal way (Zakay, 1992, Block and Zakay, 1996). *Temporal uncertainty* refers to how well the subject can estimate the duration of the task given previous experiences (Zakay, 1992, Block and Zakay, 1996). *Affective elements* represent emotional levels of the individuals while performing a task (Langer et al., 1961, Thayer and Schiff, 1975, Angrilli et al., 1997). *Arousal* refers to a state of physical activation (Schachter and Singer, 1962, Fox et al., 1967, Tipples, 2010). For example, subjects under the influence of drugs may overestimate the duration of time in comparison to others without such influence. This category is ignored in this research as none of the questions of the surveys are a good match for it. *Task complexity* refers to the effort and the characteristics of the task (Thomas and Weaver, 1975). *Temporal expectancies* refer to the accumulated previous experiences that allow the subject to generate an estimate of the duration of time for a task (Jones and Boltz, 1989, Boltz, 1993). *Absorption and attentional deployment* refer to the focus of subjects and their understanding of the task that must be performed (Tellegen and Atkinson, 1974, Glicksohn and Pavell, 1992).

In transportation research, it is likely that each of these categories are related to attributes of the travelers' environment, and to attributes of the travelers themselves. During their spatial movement, it is plausible that travelers are undergoing different levels of effort, complexity, and other factors discussed found to alter the perception of the duration of time. These categories (see figure 1.2) are *conjectured* to be more consolidated in a transportation environment as follows:

1. **Temporal Relevance**: typical commuters especially are bound to workplace time restrictions, and thus may consider choosing an optimal departure time, and route as very significant in order to respect these restrictions. Thus, it is plausible that there is a high incentive to not "waste" time during their trip to work. Furthermore, this indicates that trip purpose is significant to travelers. This research focuses only on commute trips (home to work, and work to home).

2. **Temporal Uncertainty and Temporal Expectancies**: psychologists define uncertainty in a dissimilar way to transportation researchers. The reason is that in transportation networks there's high interaction between travelers and also with elements (e.g. traffic signals, ramp meters) of the transportation system. Such interaction is likely to not be present in the tasks subjects must perform in psychological studies. On commute trips, there is likely to be a link between the temporal expectancy and the temporal uncertainty. The uncertainty of the trip will lead to additions or subtractions of the time elapsed in the trip due to those factors outside of the travelers control, and depending on the magnitude it is likely that those factors will enter into the memory of travelers. Thus, travelers that expect a trip of 20 minutes, and experience a trip of 40 minutes may tend to expect higher trips on the same chosen routes.

3. **Task Complexity and Absorption and Attentional Deployment**: travelers must experience different changes in the network during their trip. There may be: hierarchical discontinuities (e.g. moving from an arterial to a freeway); the network may exhibit circuity (i.e. travelers do not just follow a straight path towards their destination); the distance traveled of the trip in a specific hierarchy (e.g. travelers may travel mostly on freeways); and others. Thus, travelers during their trip are

likely to be stimulated by different changes that may require effort on their part to guarantee that they will arrive at the desired destination.

4. **Affective Elements**: The emotional value of the trip is also important. Travelers that may find themselves in stressful or unpleasant situations while driving may have their perception of time affected.

Moreover, the development of these *conjectured* categories is based on linking the previously discussed research on perception of time by psychologists with an existing data source (Zhu, 2010, Carrion and Levinson, 2012a). However, this is not ideal, because the study behind the data source is not specifically designed to address any research questions with regards to the perception of travel time. Thus, variables collected from the data (see chapter 4 for details) are difficult to group specifically into any of the main factors, and are grouped into a less detailed set of *conjectured* categories.

Furthermore, these *perception errors* of travel time are incorporated in a general and common decision-making process of transportation as shown in figure 1.3. In transportation, examples of these decision-making processes are mode choice, route choice, destination choice, and others. The decision-maker may only choose one alternative out of $\mathcal{J}$ possible options, and the alternatives are mutually exclusive. Each of the alternatives have sets of measured attributes, and the decision-maker knows of these attributes. The variable ξ represents the *perception errors*, and thus it is clear that the attributes themselves are connected to the *perception errors*. This emphasizes that the information as it is available in the environment still needs to be evaluated by the travelers. This mechanism also is likely to be interrelated as it evaluates every alternative. For example, the same attributes that influence the perception of travel time for a traveler are likely to be similar across alternatives. Moreover, a difficulty that is not outlined is that the same factors leading to *perception errors* of travel time may also inform the decisions of the subjects. Also, travel time is not the only attribute that may be influenced by *perception errors*.

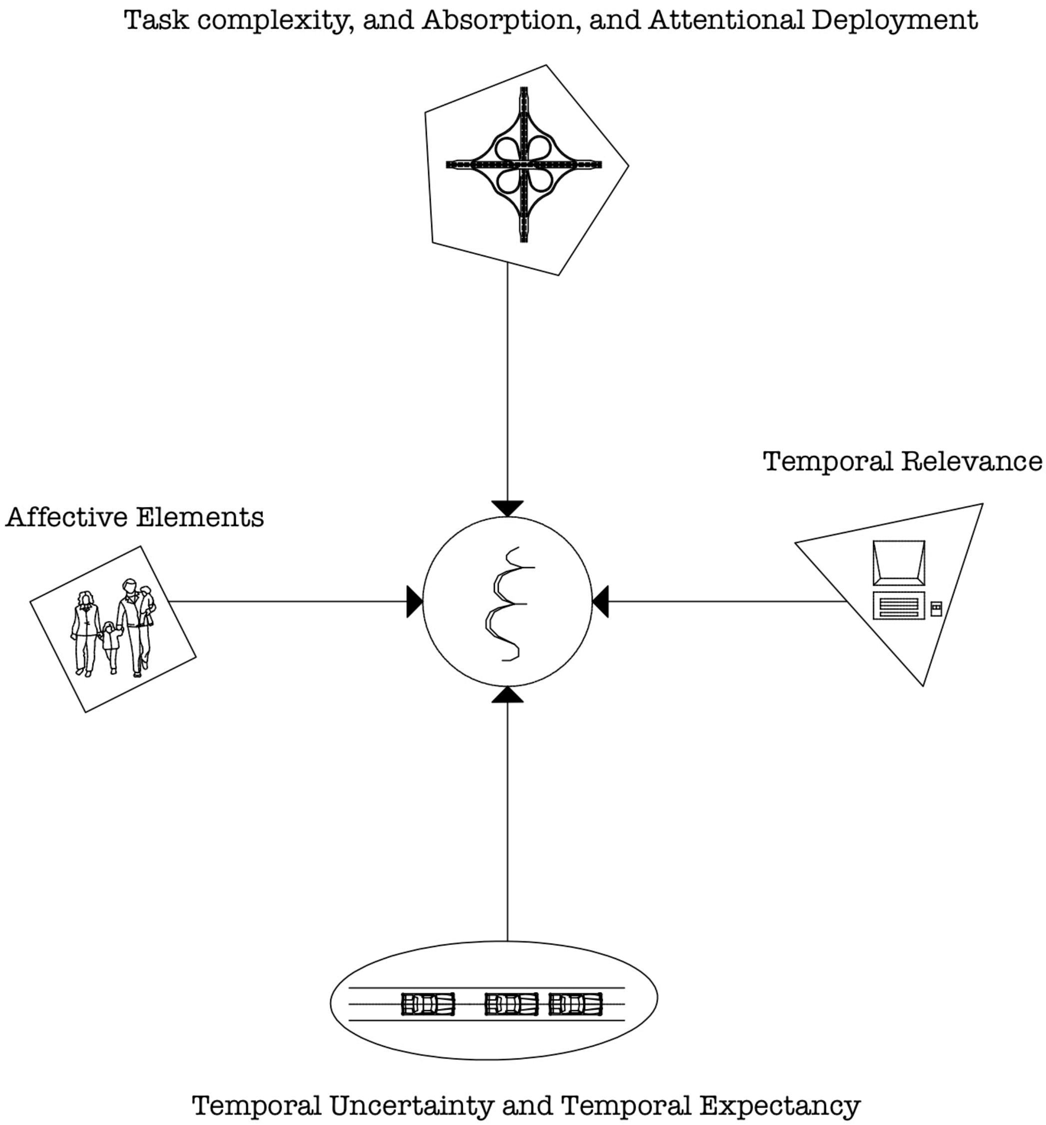


Figure 1.2: Four groups of factors that influence perception errors (ξ) of travel time in travelers.

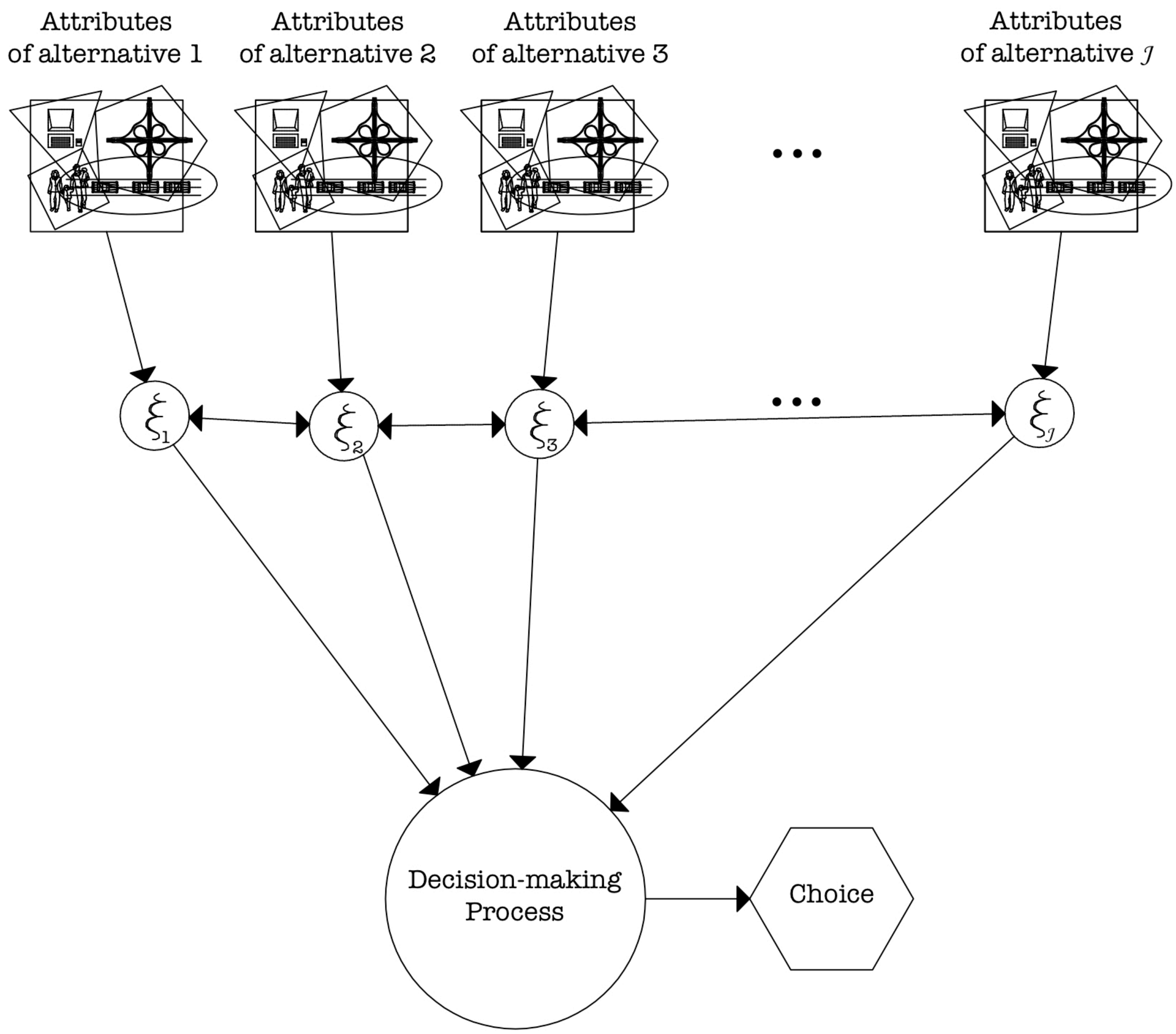


Figure 1.3: A decision-making diagram including perception errors (ξ) of travel time

1.2 Research contribution

This dissertation contributes novel results to the following areas of transportation research: travel time perception; valuation of travel time; and route choice modeling. This study presents a systematic identification of factors that are connected to the perception errors of travel time. These factors are related to time perception research in psychology through four plausible categories: temporal relevance; temporal uncertainty, and temporal expectancies; task complexity, and absorption, and attentional deployment; and affective elements. Furthermore, this dissertation provides the first empirical results that highlight the influence of perception errors in the valuation of travel time, and in the dynamic behavior of travelers' route choices. Last, but not least important, this dissertation presents the most comprehensive literature review of the value of travel time reliability written to date.

The chapters in the dissertation are organized as follows:

- **Chapter 2** reviews the literature in the three main research areas: valuation of travel time; route choice modeling; and perception of travel time. Fragments of this chapter are published in Carrion and Levinson (2012c).

- **Chapter 3** describes the data used for this analysis. It also includes descriptive statistics, and descriptive analysis of the data.

- **Chapter 4** covers the factors influencing the perception of travel time. It also describes the econometric machinery used (regression analysis), and the associated results (magnitude and direction of the influence of each factor).

- **Chapter 5** presents evidence of the influence of perception errors of travel time in the valuation of travel time. This evidence is ascertained through econometric modeling (discrete choice analysis consistent with Random Utility Theory) by systematic comparison of choices of subjects with self reported travel times, and with measured travel times from GPS devices.

- **Chapter 6** contributes empirical evidence through econometric modeling (duration analysis) of travelers reacting to day-to-day travel times on a specific route according to their own thresholds. These thresholds help discriminate whether a travel time is within an acceptable margin or not, and travelers may decide to abandon the chosen route depending on the frequency of travel times within acceptable margins.

- **Chapter 7** concludes with key findings from the dissertation.

Chapter 2

Literature Review

The literature review for this thesis encompasses two main areas: travelers' perception of travel time; and travelers' route choice behavior with a greater emphasis on the valuation of travel time reliability. There are already several studies focusing on each of these areas separately, and thus providing a comprehensive review is a difficult task. This review is selective, and only relevant studies are covered. References to further readings are provided.

2.1 Introduction

In transportation, the route choice behavior of travelers is an important pillar on which travel demand models are built. Route choice is a common decision-making process, where a traveler chooses a path connecting any two nodes from several other known alternative paths. This choice behavior is influenced by characteristics from both the traveler and the physical environment. In principle, the travelers learn about attributes of the physical environment (transportation network infrastructure and/or the set of places), and they extract the relevant information according to their own criteria. This information about the physical environment is processed also according to their own criteria, and the travelers are able to exercise their final decisions. This decision-making process is dynamic, and it is further influenced by past experience related to travelers' previous decisions (Bovy and Stern, 1990).

Typically, transportation research in route choice behavior focuses on three categories: traveler's knowledge of alternative routes; decision processes of travelers; and the influence of attributes of the traveler-road network system in travelers' route preferences. The first consists of analyzing the criteria that travelers adopt to include routes in their set of possible routes. The second focuses on the rules followed by the travelers to select their final decisions. The third examines the effects of attributes in travelers' route preferences (Ben-Akiva et al., 1984).

Most of the early research found travel time and travel distance as the main explanatory attributes for travelers' route preferences (Trueblood, 1952, Michaels, 1966, Kansky, 1967, Haefner and Dickinson, 1974, Hamerslag, 1981, Vaziri and Lam, 1983). However, research has shown that route choice behavior is not entirely encapsulated by travel time and travel distance. Other factors are also linked to the explanation of this phenomenon. These other factors include but are not limited to: travel time variability/reliability (Small et al., 2005, 2006, Tilahun and Levinson, 2010, Carrion and Levinson, 2012b); travel cost (Small et al., 2005, 2006, Tilahun and Levinson, 2010, Carrion and Levinson, 2012b); aesthetics of scenery (Zhang and Levinson, 2008); traffic information (Abdel-Aty et al., 1997, Zhang and Levinson, 2008); and others (Pal, 2004).

Studies of route choice behavior are also further categorized according to the nature of the data, and the data collection techniques employed. Generally, two data sources are used: stated preference (quasi laboratory experiments), and revealed preference (field observations of actual trips). In addition, the data collection techniques are centered around these two types of data. Techniques for stated preference include: paper-based questionnaires (Jackson and Jucker, 1982, Khattak et al., 1993, Abdel-Aty et al., 1997, Pal, 2004, Tilahun and Levinson, 2009); questionnaires with visual aids (Bartram, 1980, Goldin and Thorndyke, 1982, Tilahun and Levinson, 2010); questionnaires with simulated reality (Blaauw, 1982, Scott, 1985, Leiser and Stern, 1988, Mahmassani and Herman, 1989, Godley et al., 2002, Levinson et al., 2004, 2006); and questionnaires after designed field experiments (Zhang and Levinson, 2008). Techniques for revealed preference include: interviews in person or through the phone (Small et al., 2005, 2006); self completion questionnaires (D'Este, 1986, Duffell and Kalombaris, 1988); stalking/shadowing (Chang and Herman, 1978); and GPS tracking devices (Li et al., 2004, Li, 2004, Li et al., 2005, Carrion and Levinson, 2012a). Furthermore, the main characteristics that are variable across techniques are: cost and resources; realism and validity; degree of control of the researcher over the experiment; researcher's ability to monitor the experiment; and degree of difficulty of separating a variable's effects from others. Thus, each of the studies are trading off between a variety of characteristics (e.g. lower cost, but less realistic) when using a specific technique to collect data.

2.2 Travel time reliability

2.2.1 Concepts

The concept of travel time can be defined as the time elapsed when a traveler moves between two (distinct) spatial positions. Certainly, this definition is applicable to any transportation mode (or combinations of them) regardless of the inherent differences across them. This is expected as travel time is typically understood as a one dimensional quantity (variable). Furthermore, travel time can be divided into several components depending on the analyst. For example, travel time of public transit modes tends to be split into waiting time, in-vehicle time, transfer time, and others.

In road networks, travel time may be split into two components: free flow time, and additional time. The former refers to the amount of time it takes drivers to arrive at their destination without encountering any (or very little) traffic. The latter refers to each increase of travel time due to variations in the traffic conditions. These variations may be *predictable* (e.g. peak-hour congestion), or *unpredictable* (e.g. vehicular crashes).

The predictable variations are events (i.e. traffic congestion) expected by travelers, and thus travelers (in principle) perform the necessary adjustments to offset the added costs (e.g. departing earlier to avoid arriving late at work). Such events (i.e. traffic congestion) are by themselves a topic of interest to many researchers focusing on traffic flow theory (Daganzo, 2007). In transportation research, the morning peak-hour congestion is considered as a classic problem of trip scheduling under deterministic traffic conditions. Vickrey (1969) presented a solution to the problem with a single deterministic bottleneck model between an origin and destination, fixed and homogeneous travel demand, and endogenous departure time (i.e. trip scheduling choices). This model was further extended by Daganzo (1985), Newell (1987), Arnott et al. (1990), Arnott et al. (1993), Arnott et al. (1994), Laih (1994), Daganzo (1995), Garcia (1999), Daganzo and Garcia (2000), and others.

The unpredictable variations are directly linked to the uncertainty of travel time. This uncertainty has been divided in three elements by Wong and Sussman (1973): variation between seasons and days of the week; variation by changes in travel conditions because of weather and crashes or incidents; and variations attributed to each travelers' perception. Nicholson and Du (1997) lists also the components of uncertainty as variations in the link flows and variations in the capacity. Therefore, the unpredictable variations trace their source at both the demand side (e.g. traveler's heterogeneous behavior) and supply side (e.g. traffic signal failure) of a transportation system.

Travel time reliability is closely linked to the unpredictable variations. This suggests that travelers choose under an uncertain environment as they may fail to predict their exact travel time before scheduling their trips (i.e. choosing a departure time). In the case of predictable variations, the travelers may adjust their departure time choice, and still be certain of arriving on time at their destinations. This is true even in a transportation system with high congestion. Notice that travelers are choosing under a certain environment. Therefore, it'll be incorrect to consider predictable variations as examples of travel time (un)reliability (Bates et al., 2001). It should be noted that this travel time uncertainty may also extend to other choice dimensions (e.g. mode, route). Furthermore, the concept of travel time (un)reliability is defined as interchangeable with travel time variability (or unpredictable variation) in the transportation research literature; high variability means high unreliability, and vice versa. Consequently, it is natural to think of travel time in two dimensions: frequency and magnitude. In other words, travel time defined as a distribution in the probability theory sense. In this way, travel time (un)reliability can be associated with a measure of spread to the travel time distribution. Distinct approaches have been proposed to model travel time reliability, and they are reviewed in the subsequent section. Moreover, similarities (i.e. travel time composed of deterministic and random elements) may be drawn to other transportation modes despite the fact that the concepts were mostly explained with a focus on road transportation.

2.3 Travel time reliability and Route choice behavior

Most empirical studies of route choice in the transportation literature have focused on the estimation of the *value of travel time reliability*. This value refers to the marginal rate of substitution between travel cost, and increases in the reliability (i.e. reducing the variability) of travel time. The dominant method for the estimation of the *value of travel time reliability* is discrete choice analysis typically within the *Random Utility* framework (Ben-Akiva and Lerman, 1985, Train, 2009, Ortuzar and Willumsen, 2011). The estimation has mostly been done using stated preference data from hypothetical choice experiments. These choice experiments present scenarios with myriad of presentations to travelers. The main concern with these scenarios is whether the subjects understand the representations of travel time variability being presented to them. A secondary concern is whether subjects can perceive the situation on the hypothetical experiment as similar to actual experiences in the actual transportation network. There are few studies using revealed preference data because of few examples of experimental settings with significant travel time variation across at least two alternatives (e.g. high occupancy toll lanes); difficulties with measuring travel time data; costs associated with planning (e.g. methodology of experiment) and deployment (e.g. surveys, devices to measure travel time) of revealed preference studies; and others. Furthermore, the stated choice experiments are far more common than collected revealed preference observations for the measurement of values of travel time reliability. In addition, revealed preference studies are using measured travel time distributions as obtained from a device (e.g. loop detector, GPS device). Thus, the perception error of travelers with regards to travel time has been largely ignored.

Travel time reliability has been incorporated in route choice studies as different measures of variability of the travel time distribution. These measures are generally centered on two theoretical frameworks: Centrality-Dispersion (Jackson and Jucker, 1982); and Scheduling delays under uncertainty (Small, 1982, Noland and Small, 1995). The first is based on the idea that the travel time unreliability (or variability) is concentrated in a statistical measure of the dispersion of the travel time distribution. The second assumes that travelers have a specified time of arrival, and any *expected* late arrivals or *expected* early arrivals incurs disutilities. These disutilities are asymmetric in contrast to the Centrality-Dispersion framework that assumes all disutilities (due to unreliability) are weighted equally. It should be noted that *expected* refers to the first statistical moment of schedule delays due to late arrivals or early arrivals over the travel time distribution.

2.4 Theoretical Frameworks

2.4.1 Centrality-Dispersion

The approach is mostly known in the context of risk-return models in finance. A decision-maker looks to maximize the option's return while minimizing its associated risk. The option's return is represented by the expected value, and the risk by the variance (Markowitz, 1999). In a transportation context, the framework is based on the notion that not only travel time is a source of disutility, but also travel time variability (or unreliability). Thus, the formulation (with a linear-additive form) of the model, in a consumer theory background, is as follows:

$$U = \gamma_1 \mu_T + \gamma_2 \sigma_T \tag{2.1}$$

The traveler is minimizing the sum of the two terms (objective function for an unspecified choice dimension): the "expected" travel time of the trip, and the travel time variability of the trip. The "expected" travel time (μ_T) is included as a centrality measure (e.g. mean) of the travel time distribution. The travel time variability (σ_T) is included as dispersion measure (e.g. standard deviation) of the travel time distribution. The γ coefficients are exogenous parameters. Typically, the choice dimension is route choice, and the centrality (dispersion) measure is mean (variance or standard deviation) among studies using this approach. Mean-variance is also the usual name the approach is known in the transportation literature, despite the fact that the centrality and dispersion measures vary among studies. In the transportation literature, the framework was introduced by (Jackson and Jucker, 1982). Their original formulation is:

$$\begin{aligned} \text{Minimize} \quad & E(T_p) + \lambda_k Var(T_p) \\ p \in P_{AB} & \\ \lambda_k > 0 & \end{aligned} \tag{2.2}$$

A traveler k has a priori information of the mean $[E(T_p)]$, and variance $[Var(T_p)]$ of the travel time distribution for each route in their choice set (P) between an origin-destination pair (AB). λ_k indicates the degree of risk aversion of the traveler k. The choice dimension is the route. Succinctly, a traveler k, with a degree of risk aversion λ_k, chooses the route that minimizes the objective function (2.2) given the expected and variance of the travel time distribution. The model proposed by Jackson and Jucker (1982) is usually estimated using discrete choice methods with the linear-additive specification given in equation (2.1). In this utility form plus a travel cost variable $(\gamma_3 C)$, marginals rate of substitution may be computed to obtain important quantities such as the value of travel time (VOT), value of travel time reliability (VOR), and the reliability ratio (RR). These are defined formally in the previous order as,

$$VOT = \frac{\partial U/\partial \mu_T}{\partial U/\partial C} \tag{2.3}$$

$$VOR = \frac{\partial U/\partial \sigma_T}{\partial U/\partial C} \tag{2.4}$$

$$RR = \frac{\partial U/\partial \sigma_T}{\partial U/\partial \mu_T} = \frac{VOR}{VOT} \tag{2.5}$$

In essence, this framework is based on *expected utility theory* (Von Neumann and Morgenstern, 1944). The theory prescribes a set of axioms about how decision-makers deal with risky prospects (set of alternatives where a choice is selected) based on distinct states of nature (or states of the world). In simple words, there are several alternatives with several possible states of natures (the distribution of outcomes for each alternative), and associated with each combination of alternative and state of nature there is an outcome. In the transportation context, the set of alternatives could be routes, modes, schedules. The states of nature could be traffic signal failure, crashes, and others. The outcomes are likely to be the distribution of travel times for each alternative. In addition, the decision-maker ranks the risky prospects through the assumption of the existence of an ordinal utility function (i.e. $U = f(outcome)$; the utility function associates a single real number to each outcome), and prefers the alternative with the highest expected utility ($E(U)$). Furthermore, an important feature of the expected utility framework is based on decision-making under risk. In other words, there's a difference between risk, where probabilities are known or at least knowable, and uncertainty, where probabilities are unknown. This difference may not be particularly useful for most practical purposes, or it may be irrelevant by considering subjective probability, and the axiomatic approach of expected utility theory (Takayama, 1993). Readers may also refer to Varian (1978) and Mas-Colell et al. (1995) for treatments of expected utility theory. *Prospect theory* (Kahneman and Tversky, 1979, Tversky and Kahneman, 1992) is also another decision-making framework that is receiving attention from transportation researchers, although not as widely as *expected utility theory*. It is based on experiments (Kahneman and Tversky, 1973, Tversky and Kahneman, 1974) of decision-makers under different choice scenarios, and the common heuristics and biases dominating the decision-making process are identified.

The functional form of the utility function is not restricted by the axioms. In fact, the functional form chosen should be based on its close description of a decision-maker's behavior. The functional form determines the risk preferences of the decision-maker. Functional forms may be selected based on regression analysis of experiments (e.g. gambling games that provide observations revealing the utility function) or computationally convenient forms (Hazell and Norton, 1986).

In the transportation literature, several functional forms have been considered to understand the risk behavior of travelers. Polak (1987) considered an alternative formulation to Jackson and Jucker (1982), where he defined the utility function of the traveler as a polynomial of second degree with respect to the travel time variable (T). Formally,

$$U = \gamma_1 T + \gamma_2 T^2 \tag{2.6}$$

This functional form (2.6) is known in the microeconomics literature (Varian, 1978) as equivalent to the mean-variance model under expected utility theory. This can be seen by applying the expectation operator to (2.6), and using a simple identity ($Var(X) \equiv E(X^2) - (E(X))^2$),

$$E(U) = \gamma_1 E(T) + \gamma_2 (E(T))^2 + \gamma_2 Var(T) \tag{2.7}$$

An important consideration is that the omission of the additional term $[(E(T))^2]$ in (2.7) might bias the estimates of γ_2, especially when the formulation in (2.6) is accurate. In addition, the γ_2 indicates whether the traveler prefers alternatives (e.g. routes) with high variance of travel time (risk prone), low variance of travel time (risk averse) or only cares about the expected travel time (risk neutral). Furthermore, higher degrees of polynomials may be specified, and consequently in expected utility forms will lead to higher moments to be included. Another formulation proposed by Polak (1987) is

$$U = -e^{\gamma_1 T} \tag{2.8}$$

This functional form (2.8) is also known in the microeconomics literature (Varian, 1978); it describes a traveler with absolute risk aversion.

Senna (1994) introduced a more general form based on the previous mentioned work, where the utility function is given by a algebraic term of degree β.

$$U = \gamma_1 T^{\beta} \tag{2.9}$$

The utility function can be written in terms of expected utility, by applying the expectation operator, and by considering some simple identities such as the definition of covariance, in the following form:

$$E(U) = \gamma_1 (E(T^{\frac{\beta}{2}}))^2 + \gamma_1 Var(T^{\frac{\beta}{2}}) \tag{2.10}$$

The equation (2.10) exhibits certain properties. The β parameter estimates the degree of risk aversion/proneness by the travelers. Another important property is that the value of time and the value of variability (reliability) depend directly on the travel time distribution.

It should be noted that all the previous γ coefficients are parameters to be estimated, and expected to be negative.

2.4.2 Scheduling Delays

Historically, this approach has been linked to the departure time choice (or trip scheduling) studies. The basis for the approach rests on the time constraints (e.g. work start time) a traveler may face, and thus it associated with costs due to early or late arrival. This leads to the idea of a traveler intrinsic choice of a preferred arrival time (PAT); the point of reference that delimits whether an arrival is early or late. Gaver (1968) is one of its earliest proponents. He introduced a theoretical framework for describing variability in trip-scheduling decisions. He considered distinct head start strategies for given delay distributions along with the costs of arriving early or late. In addition, statistical estimation procedures (non-parametric and parametric) are provided to estimate the probability density distribution of the trip delay, when it is unknown to the researcher. Vickrey (1969), as described previously, also considered the trade off travelers face between queue delay, and schedule delay of arriving early or late at work. Similarly, Knight (1974) suggested that travelers consider a slack time (i.e. safety margin) between their (average) arrival time and their work start time. This safety margin allows the reduction of the probability of late arrivals, and implies that travelers have a preference of arriving early to work (i.e. existence of positive utility for the time spent at work before work start time). In essence, Knight (1974) hypothesizes that the departure time chosen happens when the marginal utility of time spent at home is equal to the marginal utility of arriving early to work plus the marginal utility of arriving late to work. Pells (1987) further argued that two opposite existing factors are at play: the need to minimize the frequency of late arrivals, and maximize the time spent at home relative to the early time spent at work. Travelers meet the first factor by allocating a safety margin, and they meet the second factor by maintaining the safety margin at required levels (i.e. safety margins are acceptable when there's more time spent at home relative to early time spent at work).

Small (1982) formulates a theoretical model based on the traditional utility maximization framework with insights from time allocation models (Becker, 1965, DeSerpa, 1971, Bruzelius, 1979). Small (1982)'s model consists of tying explicitly the departure time choice, and also adding a workplace constraint (i.e. an equation linking departure time, and working hours with merits or penalties to the wage rate; workplace policies where pay is docked by tardiness or bonuses are given for arrival on time) to the utility function of a traveler. In this way, the traveler's utility is influenced by the departure time, and also the value of time is influenced by the workplace constraint. Furthermore, he specifies a functional form for the (indirect) utility of scheduling:

$$U(t_d; PAT) = \gamma_1 T + \gamma_2 SDE + \gamma_3 SDL + \gamma_4 DL \tag{2.11}$$

This is a linear-additive form, where the γ coefficients are parameters to be estimated, and expected to be negative. In this equation, the travel time (T) is not only included but also the scheduling delays which are divided by early (SDE; defined as $Max(0, PAT - [T + t_d]))$ and late (SDL; defined as $Max(0, [T + t_d]) - PAT)$ arrivals according to a preferred arrival time (PAT), and a binary term DL to indicate whether it is a late arrival or not ($SDL > 0$). The SDE, SDL and DL terms represent scheduling considerations for the workplace constraint. t_d is the decision variable (usually a continuous real variable for mathematical models), and it represents the traveler's departure time choice. Up until this point, the scheduling delay framework describes travelers' choices under certainty. Moreover, Bates et al. (2001) points out that capacity restrictions (i.e. t_d is no longer independent of T; travelers cannot choose the same t_d as queueing is now present) of the transportation facility readily translates this framework to one extensively studied using bottleneck models (Arnott et al., 1990, 1993, 1994, Laih, 1994). This implies (as discussed in section 2.2.1) the decomposition of travel time into: free flow travel time, and additional travel time due to recurrent congestion.

The model proposed by Small (1982) is usually estimated using discrete choice methods (i.e. the departure times are discrete intervals if scheduling is the choice situation) with the linear-additive specification given in equation (2.11). In this utility form plus a travel cost variable ($\gamma_5 C$), marginals rate of substitution may be computed to obtain important quantities such as the value of travel time (VOT), value of scheduling delay early (VSDE), and the value of scheduling delay late (VSDL). Researchers often discard the lateness penalty variable (DL), because it adds a discontinuity that is inconvenient to mathematical optimization models (gradient-based), and a missing lateness penalty may translate into a higher lateness scheduling delay in econometric models. These are defined formally in the previous order as,

$$VOT = \frac{\partial U / \partial T}{\partial U / \partial C} \tag{2.12}$$

$$VSDE = \frac{\partial U / \partial SDE}{\partial U / \partial C} \tag{2.13}$$

$$VSDL = \frac{\partial U / \partial SDL}{\partial U / \partial C} \tag{2.14}$$

Scheduling Delays + Dispersion

In Noland and Small (1995), the previous scheduling approach is extended to include explicitly the uncertainty of travel time (i.e. unpredictable variation; see section 2.2.1). This uncertainty is expressed in the form of a random variable (T_r; preserving Noland and Small

(1995) notation) with a given probability density function, and with the restriction of being greater or equal to zero.

$$E(U(t_d)) = \int_0^\infty U(t_d) f(T_r) dT_r = \gamma_1 E(T) + \gamma_2 E(SDE) + \gamma_3 E(SDL) + \gamma_4 P_L \qquad (2.15)$$

The objective function of the traveler changes (also the utility function is traded for a trip cost form in Noland and Small (1995), but we choose to keep it for coherency), and now the consumer maximizes the expected utility $E(U(t_d))$ by choosing the optimal t_d (see equation (2.15)) for a given probability density function of T_r. The elements of (2.15) include the scheduling costs for early (SDE) vs. late (SDL) arrival at work presented earlier (see (2.11)), but also the last term employs the distribution of the random variable (T_r) in order to compute the probability of being late. P_L is simply $E(DL)$ (note DL is an indicator function) conditional on t_d. Therefore, the last term P_L also contains the costs of travel time unreliability as the dispersion (or variability) of the travel time distribution affects the calculated probabilities. In addition, travel time dispersion (or variability) may increase the propensity of early arrivals, and thus high earliness costs can be incurred. This implies variability and scheduling costs are related. In fact, Bates et al. (2001) argues that $\gamma_2 E(SDE) + \gamma_3 E(SDL)$ may approximate the $\gamma_2' \sigma_T$ in the centrality-dispersion model (see section 2.4.1) under certain conditions: travel time distribution is independent of departure time; $\gamma_4 = 0$ in equation (2.15) or no lateness penalty; departure time is continuous; congestion dynamics are neglected as in travel time is independent of departure time.

Recent work by Fosgerau and Karlstrom (2010) proved mathematically that scheduling models are approximately equal to mean-variance models. They indicate this can be achieved with knowledge not only of the estimated parameters (γ_1 and γ_2 in equation (2.1)) for the expected travel time and variance, but also the travel time distribution, and the optimal probability of being late. This proof follows the assumptions presented earlier by Bates et al. (2001), and also the obvious assumptions the mean of random variable (T_r; they actually use a standardized form with mean 0 and variance 1) is defined (i.e exists), and that it has an invertible distribution. These assumptions are more general than the previous ones of assuming the density function of the random variable (T_r) follows an uniform or exponential distribution (Noland and Small, 1995, Polak, 1996, Bates et al., 2001, Small and Verhoef, 2007). The interested reader should refer directly to Fosgerau and Karlstrom (2010), especially appendix A for more details. An empirical verification is also included in the paper.

Other recent work has followed different paths: inclusion of risk attitudes in scheduling models (Senbil and Kitamura, 2004, Michea and Polak, 2006, Schwanen and Ettema, 2009, Li et al., 2012); alternative formulation of schedule early (SDE), and schedule late (SDL) (Tilahun and Levinson, 2010); and scheduling preferences with non-constant marginal utilities or time-varying parameters (Tseng and Verhoef, 2008, Fosgerau and Engelson, 2011, Jenelius et al., 2011).

In Li et al. (2012), a non-linear utility specification (they assume a utility function of the form $U = \gamma_1 \frac{x^{1-\alpha}}{1-\alpha}$, where x is any variable in the model and α represents risk attitude) is used like in the other mentioned studies (Michea and Polak, 2006), but the parameter indicating risk attitude (α) was assumed random, and thus the parameters of its population density function can be estimated using a mixed logit formulation. The idea of risk attitudes has been considered before in microeconomics, and discussed implicitly in section 2.4.1.

Tilahun and Levinson (2010) introduces a new approach for measuring SDE and SDL in equation (2.11) consisting of two moments: the first representing on average how early the traveler has arrived by using that route; and the second representing on average how late that individual arrived by using that particular route. They assume that the deviation of the two moments (average late or average early) from the most frequent experience is a representative way of getting together the possible range and frequencies experienced by the travelers. Thus, this measure may considers scheduling constraints as well, albeit not separately from (un)reliability of travel time.

The scheduling preferences models (Tseng and Verhoef, 2008, Fosgerau and Engelson, 2011, Jenelius et al., 2011) generalize Small (1982) by assuming the γ parameters (except for the binary lateness penalty, which is discarded) in equation (2.11) are time-dependent (earliness and lateness penalties vary by time of day). Tseng and Verhoef (2008) introduced the formulation based on Vickrey (1973)'s model. Fosgerau and Engelson (2011) extended the formulation to account for random travel times. Jenelius et al. (2011) considered chained trips, and thus more than one activity. In addition, Jenelius (2012) extends his previous chained trip formulation to consider random travel times.

2.4.3 Mean-Lateness

This approach is widely used in passenger rail in the UK, and it was proposed by the Association of Train Operating Companies (ATOC, ATOC). It consists of two elements under the expected utility paradigm: schedule journey time ($Sched$), and the mean lateness at destination (L). The former refers to the travel time between the actual departure time and the scheduled arrival time, and the latter refers to the mean of the lateness. The lateness is defined as the time between scheduled departure and actual departure (lateness at boarding), and time between scheduled arrival and actual arrival (lateness at destination). In the original ATOC formulation (see equation 2.16) only the mean (positive [L^+]; negative values meaning early arrivals are not considered) lateness at destination is considered, but this is expanded (see equation (2.17)) in Batley and Ibanez (2009) to include the lateness at boarding (positive [B^+]; negative values meaning early departures are not considered) as well, plus the train fare is added to calculate marginal rates of substitution between temporal quantities and travel cost (e.g. value of time [VOT]). It should be noted that Batley and Ibanez (2009) also tested the inclusion of another variable ($\gamma_5 \sigma_T$) representing the standard deviation of the in-vehicle journey time.

$$E(U) = \gamma_1 SchedT + \gamma_2 L^+ \tag{2.16}$$

$$E(U) = \gamma_1 SchedT + \gamma_2 L^+ + \gamma_3 B^+ + \gamma_4 C \tag{2.17}$$

It should be noted that all the previous γ coefficients are parameters to be estimated, and expected to be negative.

2.5 Empirical evidence

Most of the initial research hinting towards travel time reliability (or predictability) was based on questionnaires ascertaining travelers' preferences, and thus it was mainly qualitative. For example, Vaziri and Lam (1983) asked commuters to list and rank possible reasons affecting their route choice, and also write others that were not listed. The results (directly) related to reliability were: "it has fewer accidents or unexpected tie-ups" (ranked fourth); and "it has smaller variation in trip times" (ranked eight). Also Chang and Stopher (1981), indicated similar results (importance of factors related to reliability) with travel mode preferences. Furthermore, Prashker (1979) was the first to explicitly account for reliability; he included different levels of variation for variables such as in-vehicle travel time, parking search time, and bus waiting time. Moreover, the research has since moved to a quantitative state. Empirical estimates have been obtained based on statistical models (typically using discrete choice methods) of the previous theoretical frameworks. The data sources for statistical modeling are usually from: stated choice experiments (i.e. stated preference) with a variety of presentations for questionnaires; and revealed choices (i.e. revealed preference) with objective travel time distributions (i.e. travel times measured by Global Positioning System [GPS] devices, loop detectors, and others). Both data sources may be combined as well to overcome some of their own deficiencies (Louviere et al., 2000). Revealed choices may be estimated using subjective travel time distributions (i.e. travel times reported by travelers memory), but this has not be done yet. The differences between subjective travel time distributions and objective travel time distributions are likely to be based on perception errors. This discussion is summarized in Figure 2.1. These reliability issues are discussed further subsequently.

2.5.1 Stated Preference Studies

Most of the estimates of valuation of reliability have been obtained through stated choice experiments. In fact, Bates et al. (2001) argued that (at the time of publication) there were no adequate real examples at the level of detail required for ascertaining reliability

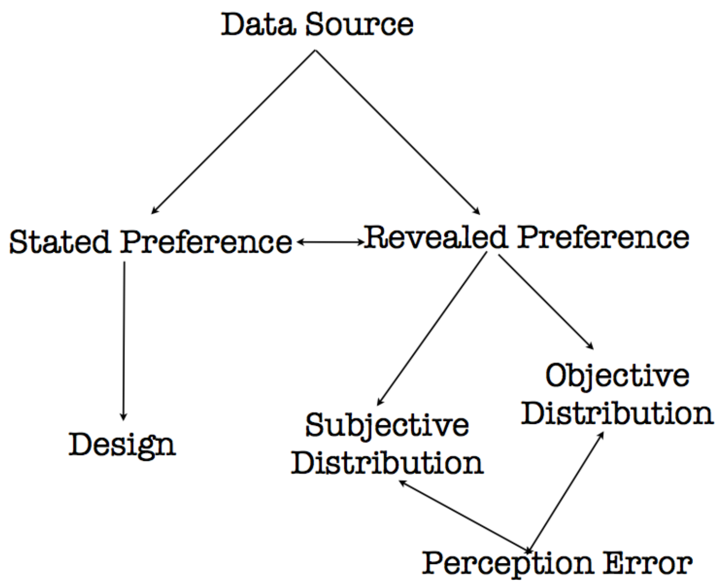


Figure 2.1: Data sources for Value of travel time reliability studies

estimates using revealed preference data (RP). Thus, they considered stated preference as the best bet, which had dominated completely the empirical studies (and its estimates) so far. However, they admitted that survey design (i.e. presentation of questions) may affect the outcome of the reliability estimates. This is likely as travel time reliability is difficult to present to subjects without any statistical background unlike travel time savings.

Early studies focused on paired comparison questions of hypothetical route alternatives. A pair was typically formed of two "usual" times and corresponding delays to one alternative of the pair. The delay was always given to the shortest "usual" time of the pair. In this way, variability measures were incorporated for the estimation of the models. Jackson and Jucker (1982) introduced the mean-variance approach (or centrality-dispersion framework covered in section 2.4.1) in order to quantify the effects of travel time (un)reliability on route choices. The analysis of the subject's stated preference was done by optimizing an objective function (a linear programming problem) in which the expectation and variance of the travel times are variables. This method also allowed for the estimation of a degree of risk aversion parameter for the subjects. Jackson and Jucker (1982) found that some commuters prefer the more reliable route, even if the expected travel time is higher in comparison to other routes with shorter expected travel time, and higher uncertainty. This result agrees with the notion of a distribution of the degrees of risk aversion in the subjects. Abdel-Aty et al. (1997) used two stated preference techniques (a computer aided telephone interview and a mail-back survey) in order to investigate the effect of travel time reliability and traffic information on commuters. The first survey consisted of offering five options, each with two routes with distinct travel times (one with the same travel time for every day, and the other with different travel times on some days) for the travelers to choose, and the second one consisted of two routes (one presumably familiar to the subjects) with similar travel time variation scheme to the previous survey, but also included a section with traffic information. The analysis of the survey data was done with binary logit models including variables such as standard deviation, mean and gender. They found that commuters consider reliability characteristics in their route choice preference, and pay attention to travel information enough to be influenced in some scenarios to deviate from their usual routes. Another finding was that males tend to choose the uncertain route more than females, and thus indicating a difference in risk attitudes related to gender.

Black and Towriss (1993) developed a different approach to the previous researchers. The approach focused more on presenting the same mean travel time for each route (i.e. alternative), but with distinct variability as it was presented with several possible travel times. In essence, the survey respondents choose between distinct options with a varying spread of travel times, mean travel times, and travel cost. Moreover, they specified and estimated linear utility function following the mean-variance approach. Their results indicated that travel time variability was a significant factor, although the magnitude was less compared to the mean travel time. In addition, they introduced the concept of reliability ratio as defined in equation (2.5). Small et al. (1999) also investigated the effects of reliability and scheduling based on Black and Towriss (1993)'s question format with minor modifications. Two alternative choices with mean travel times, a distribution of five arrival times with respect to an implied preferred arrival time, and a travel cost. Small et al. (1999) used the collected survey data to estimate mean-variance models, scheduling models, and combining both approaches (adding terms of both equations 2.1 and 2.15 plus a travel cost term ($\gamma_6 C$);

$U = \gamma_1 \mu_T + \gamma_2 \sigma_T + \gamma_3 SDE + \gamma_4 SDL + \gamma_5 DL + \gamma_6 C$) in econometric models using discrete choice methods (consistent with Random Utility models). Small et al. (1999) also further included observed heterogeneity factors interacted with travel time (mean) and travel time variability (standard deviation) variables such as: income, number of adults, number of children, and trip purpose (work trip or non-work trip). They found that survey respondents with children have a higher disutility associated with lateness compared to those without children, and also lower income respondents incur less disutility in early arrivals compared to other respondents with higher income levels. Also Small et al. (1995) and Koskenoja (1996) did an extensive exploration of the relationship between travelers occupations and other socio-demographic variables, and their preferences toward travel time reliability. Her results show differences such high income commuters with young children preferring not to increase commute time to decreased expected early arrival penalties. On the other hand, low income commuters are willing to trade 0.6 minutes of commute time to decrease 1 minute of expected early arrival penalties. In addition, Small et al. (1999) found that combining terms of the mean-variance and scheduling models lead to statistically not significant estimate of the travel time variability measure (standard deviation in this case; γ_6 is found statistically not significant from zero). This is expected as it was discussed that the mean-variance and the scheduling approaches are equivalent under certain conditions (see section 2.4.2). Furthermore, they found that nonlinearities are present in the scheduling model. This is verified by adding a quadratic term of earliness penalty ($\gamma_7 SDE^2$) that is found statistically significant different from zero. This quadratic term implies that positive utility exists up until a point of about three minutes. In addition, the nonlinearity results indicate that the penalties (early or late) are present and are related to the preferred arrival time. Small et al. (1995) and Koskenoja (1996) also explored nonlinearities in mean travel times, and early penalties using quadratic terms. The scheduling preferences models (Tseng and Verhoef, 2008, Fosgerau and Engelson, 2011, Jenelius et al., 2011, Jenelius, 2012) hypothesize that the scheduling terms ($\gamma_3 SDE + \gamma_4 SDL$; they discard DL because of its discrete nature) are actually functions that depend on the time of day. In fact, Tseng and Verhoef (2008) tested using stated preference data that the value of travel time is different across time of day. Liu et al. (2007) also indicated the non constancy of the value of travel time, and value of reliability using loop detector data. Therefore, it can be argued that nonlinearities are starting to be considered in the recent mathematical models of scheduling preferences.

In the late 1990s and 2000s, the stated preference research focused on designing better presentations of questions about variability. Cook et al. (1999) and Bates et al. (2001) asserted that the presentation of variability in the questions has a significant impact in the estimates, because of a mismatch between the respondents and analysts understanding of the abstract situation. Thus, analysts must validate the understanding of their questionnaires with the survey respondents. Bates et al. (2001) and Cook et al. (1999) verified the understanding of respondents by presenting closely matching pairs of questions. They found that about 90% of respondents correctly identified the differences in the questions, except in cases where zero delay was included, and respondents will choose the more variable (less reliable) alternative of the pair. Bates et al. (2001) and Cook et al. (1999) proposed an alternative design for the presentation of variability. This design consists of circular arrangement of arrival times with respect to a given preferred arrival time. Each arrival time is represented by a box indicating how many minutes early or late the respondent will arrive. Bates et al. (2001) and Cook et al. (1999) included education phases to increase the likelihood of survey respondents understanding of their circular presentation. Copley et al. (2002) studied different presentations (linear arrangements of possible travel times, circular arrangements of possible travel times, and histogram representation of possible travel times) of travel time variability. A qualitative approach by interviewing respondents suggested a preference for linear arrangements and histograms presentations of variability. Copley et al. (2002) prefers the histogram representation, because it can present a large volume of information, and their qualitative research showed that it was understood with little effort.

Other researchers also tried alternative presentations. Hensher (2001) used bar diagrams dividing the total travel time into: free flow, slowed down, stop/start, and uncertainty. The bars also provided numbers for the amount of minutes of each component of the total travel time for pairs of alternatives. The alternatives also included a travel cost component in order to calculate trade offs between cost and the distinct components of time. It should be noted that Hensher (2001) was more concerned with investigating the values that travelers assign to the distinct components of the total travel time rather than travel time reliability. Also, the uncertainty component is actually more closely related to the schedule delays (allocated extra time to avoid arriving late) rather than measures of the travel time variability (e.g. standard deviation). Hollander (2006) uses a very different presentation compared to the previous discussed researchers. Hollander (2006)'s survey design consists of five bars per alternative indicating the time of departure (e.g. 8:15) on the top of the bar, and the time of arrival at the bottom of the bar (e.g. 8:30). In this way, travel times are not given in terms of minutes explicitly. In addition, travelers are told the time they should be at their destinations, explicitly. Hollander (2006) estimated a scheduling model, and a mean-variance model. The results indicate that the reliability ratio was very low 0.1 (this is significantly small compared to most studies) in the mean-variance model, and most users were willing to pay more to avoid arriving late in the scheduling model. Asensio and Matas (2008) uses a similar presentation of variability as Small et al. (1999). Asensio and Matas (2008) also tests scheduling and mean-variance models. They find that the inclusion of the variability measure plus scheduling delay measures resulted in loss of statistical significance in the reliability variables of both models with the exception of schedule delay late. Thus, indicating a correlation between both approaches as theoretically expected, and already discussed. Tilahun and Levinson (2010) introduces a variability format consisting of a histogram for each alternative in a pair. They also introduce an education phase to explain to survey respondents what the histograms convey. They test a mode-variance model (mode is the most frequent travel time shown in the histograms), mode-right range (100th percentile - 50th percentile), and introduce a new measure consisting of two moments (one representing earliness, and another lateness). Tilahun and Levinson (2010) found a reliability ratio of 0.89 for the mode-variance model. They also found that survey respondents value lateness (in their proposed measure) similarly to travel time savings. Li et al. (2010) introduce two distinct questionnaires representing variability based on Hensher (2001)'s format and Small et al. (1999)'s format. The first questionnaire contains three sections: average travel time experience, probability of time of arrival, and trip costs. The first section presents a division of average travel time very similar to Hensher (2001)'s format. The second section presents the arrival time with respect to a implied preferred arrival time very similar to Small et al. (1999) distribution of arrivals. The third section includes travel costs; a running cost is presented in addition to tolls costs. The second questionnaire is similar to the first questionnaire, except that the sections are not divided, and the distribution of arrival times is replaced with a row indicating the trip time variability (i.e. amount of minutes more or less with respect to the travel time). Travel costs are presented as taxi fares, and toll costs. Li et al. (2010) tested the questionnaires with commuters and non-commuters, and found that non-commuters value less travel time savings, lateness penalties, and travel time reliability than commuters. The non-commuters' reliability ratio is higher than commuters. In addition, Li et al. (2010) argued that the survey design similar to Small et al. (1999) (first questionnaire) is better understood by survey respondents in comparison to the survey design similar to Jackson and Jucker (1982) (second questionnaire). It should be noted that there are differences between Li et al. (2010)'s second questionnaire and Jackson and Jucker (1982)'s questionnaire even though Li et al. (2010) considers them as similar. An important difference is that Jackson and Jucker (1982) presents variability as number of additional minutes of delay per week, and Li et al. (2010) presents delays by more or less minutes with respect to the travel time.

Tseng et al. (2009) is an important contribution to the design of stated preference surveys for analyzing travel time variability. They use face-to-face interviews to investigate the understanding of subjects with most of the previously discussed questionnaires (Small et al., 1999, Bates et al., 2001, Copley et al., 2002, Hollander, 2006). The analysis consisted of questions about the respondents subjective preferences with regards to the formats, and questions that tested for consistency and logic the perception of respondents with regards to reliability presented in the questionnaires. Tseng et al. (2009) found that Small et al. (1999)'s format is preferred, and understood by most of the respondents. Copley et al. (2002)'s format showed signs of difficulty in understanding the probabilities from the graph by some of the respondents. Hollander (2006)'s format received mixed results. Tseng et al. (2009) recommends not using this format. In addition, Bates et al. (2001)'s format was not preferred compared to other formats by respondents.

In summary, stated preference studies have focused on exploring distinct presentations of variability to survey respondents based on mean-variance and scheduling approaches. Unfortunately, validation and testing whether survey respondents can understand the presentation has not received enough attention, except for early studies (Cook et al., 1999, Bates et al., 2001, Copley et al., 2002), and a recent pioneer study (Tseng et al., 2009). Most researchers agree that the variability presentation by Small et al. (1999) should be the current preferred presentation of travel time variability. It has been found it is understood by survey respondents, and it can be used to estimate both mean-variance and scheduling models. An alternative to Small et al. (1999)'s format is the histogram graphical representation (Copley et al., 2002, Tilahun and Levinson, 2010). This presentation seems to be understood by respondents, but an amount of effort is required to educate the respondents. Furthermore, there's still a need to test how subjects' preferences of travel time variability in stated choice experiments (abstract situations) compare to subjects' preferences in actual observed trips (see Hensher (1994), Louviere et al. (2000), and Hensher (2010) for discussions about SP vs. RP). Figure 2.2 presents images of some of the discussed surveys.

Figure 2.2: Examples of SP experiments of travel time variability

(a) Source: Jackson and Jucker (1982)

1. Usual time: 50 minutes
   Possible delays: None
2. Usual time: 40 minutes
   Possible delays: 20 minutes once a week.

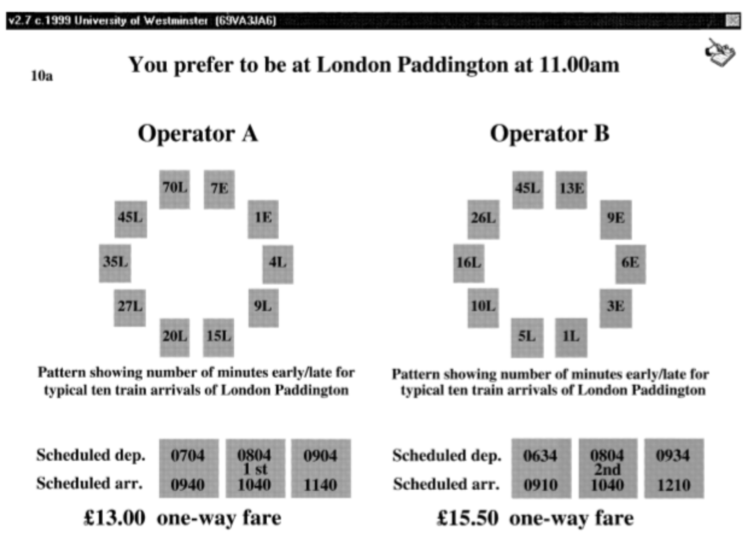
(b) Source: Small et al. (1999)

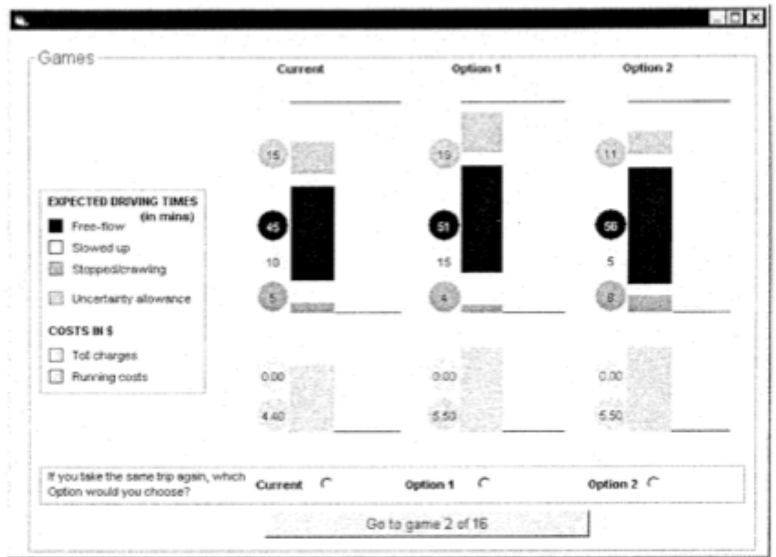
EXPERIMENT #1 (SAMPLE QUESTION)

PLEASE CIRCLE EITHER CHOICE **A** OR CHOICE **B**

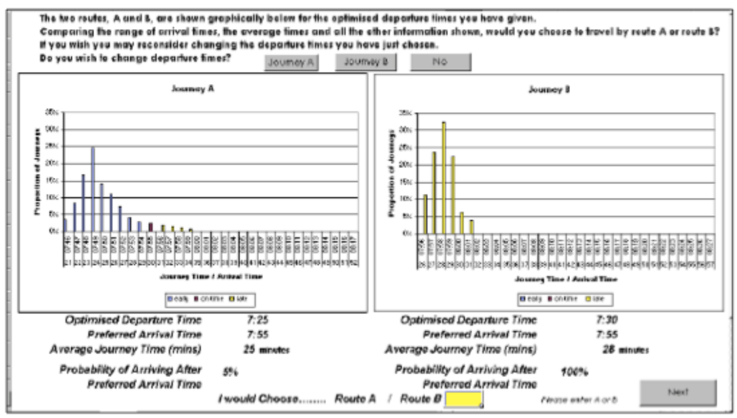
Average Travel Time: **9 minutes**	Average Travel Time: **9 minutes**
You have an equal chance of arriving at any of the following times:	You have an equal chance of arriving at any of the following times:
7 minutes early	**3 minutes early**
4 minutes early	**3 minutes early**
1 minute early	**2 minutes early**
5 minutes late	**2 minutes early**
9 minutes late	**On time**
your cost: **$0.25**	your cost: **$1.50**
Choice A	**Choice B**

(c) Source: Bates et al. (2001)



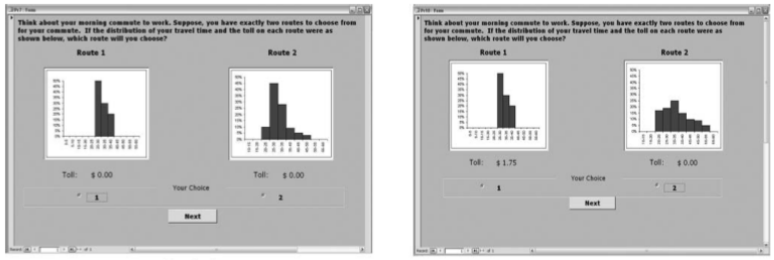


(a) Source: Hensher (2001)

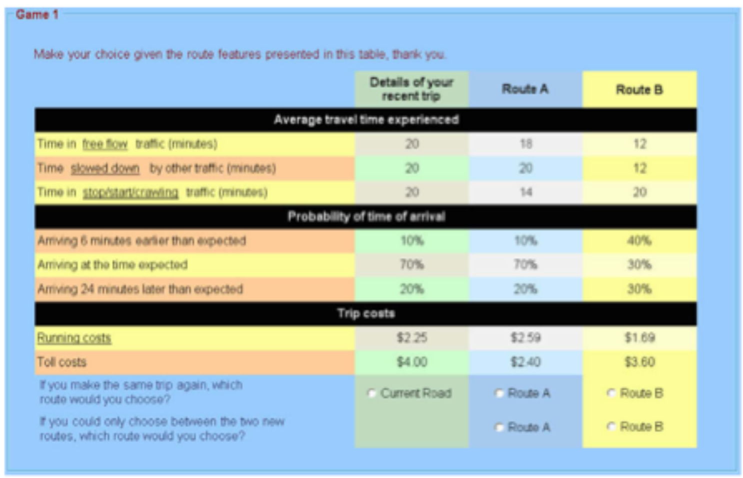


(b) Source: Copley et al. (2002)

(c) Source: Hollander (2006)



(a) Source: Tilahun and Levinson (2010)



Game 1

Make your choice given the route features presented in this table, thank you.

	Details of your recent trip	Route A	Route B
Average travel time experienced			
Time in free flow traffic (minutes)	20	18	12
Time slowed down by other traffic (minutes)	20	20	12
Time in stop/start/crawling traffic (minutes)	20	14	20
Probability of time of arrival			
Arriving 6 minutes earlier than expected	10%	10%	40%
Arriving at the time expected	70%	70%	30%
Arriving 24 minutes later than expected	20%	20%	30%
Trip costs			
Running costs	$2.25	$2.59	$1.69
Toll costs	$4.00	$2.40	$3.60
If you make the same trip again, which route would you choose?	Current Road	Route A	Route B
If you could only choose between the two new routes, which route would you choose?		Route A	Route B

(b) SP Scenario with arrival distribution. Source: Li et al. (2010)

Practice Game

Make your choice given the route features presented in this table, thank you.

	Details of your recent trip	Route A	Route B
Time in free flow traffic (minutes)	15	21	12
Time slowed down by other traffic (minutes)	10	10	8
Time in stop/start/crawling traffic (minutes)	2	2	3
Trip time variability (minutes)	+/- 8	+/- 9	+/- 8
Taxi fare	$30.70	$27.63	$18.42
Toll costs	$4.00	$0.00	$0.70
If you make the same trip again, which route would you choose?	Current Road	Route A	Route B
If you could only choose between the two new routes, which route would you choose?		Route A	Route B

(c) SP scenario with explicit variability. Source: Li et al. (2010)

2.5.2 Revealed Preference Studies

There are few studies using revealed preference data investigating travel time reliability in the transportation literature. The main reasons for their scarcity are: few examples of experimental settings with significant contrast of travel time variation across at least two alternatives (e.g. high occupancy toll lanes); difficulties with measuring travel time data; costs associated with planning (e.g. methodology of experiment) and deployment (e.g. surveys, devices to measure travel time) of revealed preference studies; and others. In addition, revealed preference studies vary by the source of travel time measurements: objective travel time distribution (measured by devices such as loop detectors); and subjective travel time distribution (travel times reported by the subjects).

Objective travel time distribution

In Small (1982), revealed preference data of trip timing (i.e. individuals were asked about their arrival times, and work-start times) of commuters in the San Francisco Bay Area was fitted to the scheduling delay with choice model where the dependent variable were twelve intervals of time of 5 minutes. The travel-time data were obtained using a road network maintained by the Metropolitan Transportation Commission, and also it was supplemented by floating car observations. Small (1982) used this data to fit his proposed scheduling model (see section 2.4.2). He found that travelers prefer early arrival, and additional travel time to late arrival, and prefer early arrival to addition travel time. In mathematical terms, the inequality $\gamma_2 > \gamma_1 > \gamma_3$ in equation (2.11). This relationship has been enforced in most (if not all) theoretical models considering cost functions (or utility functions) including early vs. late costs.

Most of the revealed preference research has been done by analyzing data collected from California State Route 91 (SR-91) in greater Los Angeles. A section of 10 miles (16 km) of this freeway includes four untolled lanes, and two high occupancy toll lanes in each direction. The high occupancy toll lanes opened in 1995, and the tolls assigned to the lanes vary by time of day. In 1997 and 1998, Lam and Small (2001) collected revealed preference data through mail surveys from drivers identified in this corridor through their license plates. Travel-time data is collected from loop detectors. Lam and Small (2001) fits mean-variance approach for route choice models to the data. They did not use toll to represent travel costs, but rather use a proxy variable representing wage rate. They also estimate the models using two distinct measures of centrality (mean and median), and two measures of dispersion (standard deviation and 90th percentile minus median). In addition, they estimate route choice and time of day models. They consider the mean-variance (only the median and 90th percentile minus median) and scheduling approach for such models. Other models considered are route and mode choice, and transponder choice as well with similar approaches. They are able to estimate the value of travel time savings, and value of reliability, but they express doubt of these estimates as they are obtained through aggregated data (loop detectors) based on many assumptions. Small et al. (2005) and Small et al. (2006) collected both RP (actual preferences of subject's lane choice) and SP (hypothetical scenarios to examine subject's lane choice) observations, and consequently enriched their statistical model by pooling both types of data of SR-91 from 1999 to 2000. The collection consisted of three surveys: the first survey was a telephone interview of actual travel (revealed preference), and the other two were mail-back questionnaires (the first one about actual travel [revealed preference], and the other one about hypothetical scenarios [stated preference]). The set of actual alternatives was composed of High-Occupancy Toll lanes (HOT) and General Purpose Lanes (GPL). Commuters using the HOT lanes require an electronic transponder to pay a toll, which varies hourly. It should also be noted carpools (High Occupancy Vehicles [HOVs]) are allowed in the HOT lanes with a discount. The set of hypothetical alternatives remained the same as the actual with the exception of changing the values of variables such as time, cost and reliability. These changes allowed for the preferences of the subjects to be inferred based on their unique pattern of responses to trade-offs among the different hypothetical scenarios. The data was analyzed by a discrete-choice model; a utility function was specified containing attributes for the alternatives including toll, travel time and reliability. This statistical model approach allows for the estimation of the well known value of time (VOT), and the value of reliability (VOR). The latter value represents the susceptibility of the commuters to (un)reliability in monetary terms, and it is calculated as the ratio between the parameters of travel reliability and travel cost (toll cost in the study). This VOR represents the marginal rate of substitution between travel cost, and travel reliability. Right ranges (80th - 50th percentiles) on the travel time savings distribution (differences between travel time distributions of GPL and HOT) are used as (un)reliability measures. Another important feature of the model is the inclusion of a carpool variable in order to control for systematic bias. However, besides all these similarities the studies differ in certain key areas. The first study Small et al. (2005) focuses solely in formulating a lane choice model (using mixed logit) by combining the RP and SP data. The results of the model indicate travel time and reliability to be significant, and that the heterogeneity in these factors is significant as well (thus implying the significance of the heterogeneity of VOT and VOR). In contrast, the second study Small et al. (2006) models not only lane choice, but also vehicle occupancy and transponder acquisition. It also extends the previous study Small et al. (2005) by using simulations to analyze distinct highway pricing policies besides the current one at CA-91. The policies simulated include: no toll, general purpose and HOV, general purpose and HOT, and combinations of the preceding cases. The objectives of these simulations are to point out the significance of the heterogeneous preferences of commuters to highway policymakers, and, as Small et al. points out, the current use of homogeneous preferences fails to account accurately for different policies working together. It should be noted that highway pricing policies are typically developed for congestion relief. The main notion being that congestion is a negative externality of the transportation system, and the use of pricing schemes will reduce avoidable trips, and persuade travelers to reconsider their activity patterns in time and space. Readers may refer also to Yan (2002) for further information. It should also be emphasized that the nature of the survey methods employed didn't allow for some of the variables to be measured during each of the subject's trips. For example, travel time was obtained by field measurements (performed by others instead of the subjects) corresponding approximately to the travel periods of the subjects. Thus, these measurements may have affected the accuracy of the data in the model. Lastly, Liu et al. (2004) also used data from SR-91, except that they used loop detector data. They propose an alternative method to RP and SP data. They consider aggregated counts from loop detectors, and origin-destinations from ramps along the freeways. Liu et al. (2004)'s VOT and VOT estimates are similar to those from the previous studies in the SR-91 freeway.

Ghosh (2001) uses data collected from the high occupancy toll (HOT) lanes in Interstate 15 in San Diego. In 1998, the tolls of the high occupancy toll lanes started being adjusted according to travel demand, and in order to maintain free flow traffic conditions. The tolls range were typically from USD$0.50 to USD$8.00 for single occupancy vehicles (i.e. solo drivers). High occupancy vehicles (i.e. carpoolers) continued to use the lanes without paying tolls. Ghosh (2001) used panel data collected by San Diego State University from 1998 to 1999. The panel consists of samples of HOT lane users (those that have transponders), other I-15 users, and users of I-8. He also used choice-based sampling to avoid over sampling HOT subscribers (transponder users). For travel time data, Ghosh (2001) asks subjects about the ramps they used to access the lanes, and uses traffic speeds from loop detector data, and estimates time savings based on arrival times from subjects. Moreover, Ghosh (2001) estimated mode choice models (choices where subscribe, nonsubscriber, carpooler, and others similar) using mean-variance approach. He considered for centrality measure the median, and for dispersion the 90th percentile minus the median. In addition, stated preference data was collected in one of the panel waves asking simple questions whether a toll value and time saved will be acceptable to the subjects. However, the SP data was only used to estimate VOT as it did not have any connection with travel time variability. Therefore, only RP data is used to estimate value of reliability using mean-variance approach. Ghosh (2001) found estimates similar (but slightly higher) than the previously discussed California studies.

Bhat and Sardesai (2006) collected revealed preference of mode choices from a web-based commuter survey in Austin, Texas. In addition, they designed a stated preference experiment where its attributes are pivoted from the revealed preference's attributes. The travel time data is based on self reported travel times, but the travel time variability is only found in the stated preference experiment. They estimate mixed logit models using mean-variance approach (centrality: mean of travel time; dispersion: standard deviation). They estimated VOT and VOR estimates for SP model, and joint SP-RP models.

Another group of researchers also studied the high occupancy toll lanes, but in Interstate 394 at Minneapolis, Minnesota. Liu et al. (2007) uses loop detector data to estimate VOT and VOR based on the method of aggregate data discussed in Liu et al. (2004), but with variations to allow the estimates to be in function with time of day. They found that VOT and VOR values varied from about USD$5 to USD$30. A more recent study by Carrion and Levinson (2012b) used Global Position System (GPS) devices and transponders, and proposed an experimental design to estimate VOT and VOR. Carrion and Levinson (2012b) considered the route choice (untolled lanes, tolled lanes, and signalized arterials parallel to the I-394 corridor) of recruited subjects from the western suburbs, and with work locations near downtown Minneapolis. Each of the subjects was equipped with a GPS device and a transponder. Thus, subscription to the HOT lanes was not an issue. Succinctly, they proposed an experimental design where subjects will drive several weeks on each route alternative, and the last two weeks will be allowed to freely choose between the alternatives. The GPS devices allowed to ascertain the RP choices, and also the travel times and other related commute level data of the subjects. Surveys were also administered to collect socio-demographic data, and also to measure travelers preferences with regards to the alternatives. Unfortunately, the study suffered from high attrition due to the requirements of the experimental design, and also experience some data lost with regards to the GPS devices. Carrion and Levinson (2012b) estimated mixed logit models using the mean-variance approach with different measures (centrality: mean, median; dispersion: standard deviation, 90th percentiles minus median, and interquartile range). The estimates of VOT and VOR were significantly lower (about USD$8) than the previous studies, but the estimates confidence intervals were wide enough to include the previous estimates of some of the other studies. In addition, Carrion and Levinson (2012a) uses GPS data from another experiment (Zhu, 2010). The GPS data was originally collected to study the travel behavior of commuters after the Interstate 35W bridge collapse in Minneapolis, Minnesota. Carrion and Levinson (2012a) used this data to fits a bridge choice model where travelers chose the new I-35W bridge or any of the possible alternatives. They included variables such as mean travel time, and standard deviation of travel time for the alternatives, and thus also calculated reliability ratios, but not VOT and VOR estimates.

In summary, high occupancy toll (HOT) lanes of SR-91, I-15 in California, and I-394 in Minnesota have become the experimental settings for RP studies. A significant problem with the RP studies is the trade-off between measuring travel time data, and the cost associated with the devices to measure such data. Loop detector data is typically collected by many department of transportation in the US, and in many cases freely available to researchers. However, loop detector data may be difficult to adapt for statistical estimation as Lam and Small (2001) noted. Another approach followed by Yan (2002) and Small et al. (2005, 2006) consisted of driving on similar time periods as the subjects and measuring travel time. This approach may approximate the actual travel time the subjects experience when they revealed their choices. Lastly, GPS devices measure very detailed commute data, and also can be used to ascertain the revealed choices of the subjects. However, it is important to cautiously design methodologies that avoid the problems Carrion and Levinson (2012b) experienced. Furthermore, the mean-variance approach dominates the RP models, because most likely preferred arrival times of the subjects were not collected.

Subjective travel time distribution

Up until this point, it has been assumed that travelers choose optimally under the objective travel time distribution (i.e. the perception error of travelers is close to zero). Bates et al. (2001) argues that it is likely travelers are optimizing according to their own divergent view of the objective distribution (i.e. based on actual measurements). Consequently, travelers will differ in their optimal solutions depending on the degree of distortion of their subjective distribution with regards to the objective distribution. This very likely as it has been shown in the transportation literature for different types of travel time such as waiting time (Levinson et al., 2004, 2006).

Recently, Peer et al. (2010) and Peer (2013) studied the travelers' perception of their morning commute. Basically, they compared reported travel times by subjects from questionnaires, and compared them to their travel times from camera data. In essence, they compared reported travel time distributions (subjective) to camera travel time distributions (objective). They found that certainly perception error is an issue that need to be taken in consideration. This result should be emphasized as more RP studies may be underestimating or overestimating the value of time, and value of reliability as the objective travel time distributions differ to subjective travel time distributions. In other words, travelers may see worthwhile savings and predictability (low variability) that do not match the actual conditions.

At the moment, no studies have been conducted with regard to the travelers' *perceived* travel times (i.e. subjective travel time distribution). Thus, it is currently unknown whether travelers' *perception error* has a significant impact on estimates of VOT and VOR. One of the main objectives of this thesis is to contribute results with regards to this unanswered research question.

2.5.3 Summary

In stated preference studies, researchers have focused in the development of choice experiments with a variety of presentations of travel time variability. The objective is to find a presentation that matches the survey respondents understanding of the abstract situation with the analysts' intentions of the abstract situation. However, most researchers have not focused on validating such understanding, and it has become difficult to ascertain which estimates are more plausible than others (especially as there are few revealed preference studies). Fortunately, some of the early studies (Cook et al., 1999, Bates et al., 2001, Copley et al., 2002) have focused on testing whether qualitatively or quantitatively the survey respondents understanding of several proposed presentations. A recent pioneer study by Tseng et al. (2009) further studied this validation concern, and found that travelers are more capable of understanding with ease the Small et al. (1999)'s format. Also, studies with histograms (Copley et al., 2002, Tilahun and Levinson, 2010) may be considered as well as long as survey respondents are educated with regards to what the histograms convey. To date, most research has not addressed another important issue of how subjects' preference of travel time variability in stated choice experiments compare to the subjects' preferences in actual observed trips.

In revealed preference studies, the literature is dominated by data from high occupancy toll lanes (SR-91 and I-15) from California (Ghosh, 2001, Lam and Small, 2001, Yan, 2002, Liu et al., 2004, Small et al., 2005, 2006). These lanes have become a experimental setting for reliability study as in some cases the contrast between high occupancy toll lanes and parallel untolled lanes in terms of travel time savings and reliability is significant. The main problem with RP studies (besides the cost of planning and deployment of such studies) is the collection of usable travel time data of the subjects. Researchers have used loop detectors, in-field measurements (driving on similar travel periods and the subjects), and GPS devices. The loop detectors require several assumptions (some questionable) and processing to estimate usable travel time data for the studies. In-field measurements may be more usable and require fewer assumptions, but they do not reflect exactly the travel times experienced by the travelers. GPS devices (Zhu, 2010, Carrion and Levinson, 2012a,b) measure very detailed commute level data of the travelers, but caution must be undertaken in methodological designs and possible requirements as attrition is a concern. In addition, the Centrality-Dispersion framework dominates RP studies. This is possible as preferred arrival times were probably not collected, and Centrality-Dispersion variables are less difficult to measure in comparison to variables required by Scheduling delay models. Moreover, there's an important gap between objective travel time (measured from devices) and subjective travel time (reported by subjects) that needs to be addressed. Subjects are likely to make their decisions based on their perceptions of travel times that should be connected to the objective distribution but with a distortion. Recent research (Peer et al., 2010, Peer, 2013) studied the travelers' perception of their commute. Basically, they compared reported travel times by subjects from questionnaires, and compared them to their travel times as measured from devices. They found that certainly perception error is an issue that need to be taken in consideration.

2.6 Perception of travel time

Psychologists have shown clear interest into the behavioral and cognitive mechanism of perception of time. They have classified the perception of time into three main categories: subjective time passage (i.e. perception of the speed that time passes); estimation of time duration; and simultaneity and succession of time. The estimation of time duration is the most frequently studied category by psychologists, and thus it is better understood. It is also the dimension of time perception that will be focus of this study, and it has been the focus of most studies investigating perception of travel time in the transportation literature. Main factors identified in the duration of time are: temporal relevance, temporal uncertainty, affective elements, arousal, task complexity, temporal expectancies, absorption and attentional deployment. *Temporal relevance* refers to the significance of time for performing a task in an optimal way (Zakay, 1992, Block and Zakay, 1996). *Temporal uncertainty* refers to how well the subject can estimate the duration of the task given previous experiences. Thus, results indicate that when a task is commonly performed, its uncertainty is low, but when a task is uncommonly performed, its uncertainty is high. In addition, tasks with high levels of relevance and uncertainty are associated with estimates of duration of time tend to be longer. In contrast, tasks with low levels of relevance and uncertainty are associated with shorter duration of time (Zakay, 1992, Block and Zakay, 1996). *Affective elements* represent emotional levels of the individuals while performing a task. For example, subjects experience fear estimate the duration of time to be shorter than those neutral (Langer et al., 1961, Thayer and Schiff, 1975, Angrilli et al., 1997). *Arousal* refers to a state of physical activation. For example, subjects under the influence of drugs may overestimate the duration of time in comparison to others without such influence (Schachter and Singer, 1962, Fox et al., 1967, Tipples, 2010). *Task complexity* refers to the effort and the characteristics of the task. Research indicates that high complexity leads to overestimation of the duration of time. In general, subjects that process more events during the time at hand will tend to overestimate as they will have more memories (Thomas and Weaver, 1975). *Temporal expectancies* refer to the accumulated previous experiences that allow the subject to generate an estimate of the duration of time for a task. Results indicate that previous durations of time will guide the duration of time for a new task (previously performed), and also update experiences (Jones and Boltz, 1989, Boltz, 1993). *Absorption and attentional deployment* refer to the focus of subjects and their understanding of the task that must be performed. Subjects that do not focus and/or do not understand how to perform the task at hand will take further time figuring the details of it, and thus may overestimate the duration of time (Tellegen and Atkinson, 1974, Glicksohn and Pavell, 1992). Readers may refer to Allan (1979) and Madalina (2011) for more details.

In the case of perception of travel time in the transportation literature, most of the studies as previously mentioned have focused on the estimation of time duration of the travelers. In essence, the travel times reported by the travelers are analyzed through several methods with the actual travel times that the subjects experienced. Transportation researchers may have control over the environment similar to psychological researchers through computer-based simulations, and/or fixed-base vehicle simulators (Levinson et al., 2006, 2004). On the other hand, transportation researchers may collect data from field observations through questionnaires, cameras, GPS devices, and others (Peer et al., 2010, Peer, 2013). It should be noted that there is an obvious trade-off between the analyst's control over the environment, and the realism of the environment to the subjects.

In the case of studies using computer-based simulations, Levinson et al. (2004, 2006) study the travelers' preferences towards waiting times during distinct traffic conditions (e.g. free flow traffic). They used computer administered stated choice experiments with written travel times and stated choice experiments based on subjects' travel times inside vehicle simulators. The results indicated that subjects perception of the travel times as presented in the computer administered experiments, and the experiments with vehicle simulators are significantly different.

In the case of field observations, Peer et al. (2010) and Peer (2013) studied the travelers' perception of their morning commute. The data sources are reported travel times by subjects from questionnaires, and travel times as observed from cameras. The reported travel time distributions are compared to the camera travel time distributions.

Recent research (Parthasarathi, 2011, Parthasarathi et al., 2012) has identified a set of factors describing the underlying structure of road networks as contributors to the travel time perception of travelers. Parthasarathi (2011) and Parthasarathi et al. (2012) used linear regression analysis on data of two sources: 2000 Twin Cities Travel Behavior Inventory (TBI); and Surveys from the I-35W Bridge collapse and reopening. The TBI is a comprehensive one day house travel survey prepared by the Metropolitan Council, and the Minnesota Department of Transportation (Mn/DOT). Participants provide a record of all trips on the surveyed day along with individual and household socio-demographic data (TBI, 2003). The surveys from the I-35W Bridge collapse and reopening refer to: two hand-out/mail-back paper surveys; one computer-based internet survey; and GPS data collected from the vehicles of subjects. The purpose of the surveys was to understand the impacts of the bridge collapse, and reopening on traveler behavior (Zhu, 2010). Furthermore, the factors used in Parthasarathi (2011) and Parthasarathi et al. (2012) are measures based on Xie and Levinson (2007) representing the hierarchical, and/or topologic features of road networks. Hierarchical attributes should be understood as those characteristics that capture the differentiation (i.e. heterogeneity) that exists in road networks. Topologic attributes are those characteristics that identify the distinct connection patterns and connectivity of different configurations of links and nodes of road networks. In addition, Parthasarathi (2011) and Parthasarathi et al. (2012) include socio-demographic variables to control for observed heterogeneity in the sample. Results indicated that the factors related measures of the structure of road networks were statistically significant, and the socio-demographic variables (e.g. income) were not.

In summary, subjects' perception of travel times has been found to be a significant factor in studies. Travelers overestimate or underestimate the actual travel times they experience. Therefore, it is likely to influence their travel decisions. Moreover, only recently the *perception error* of travelers has been connected to the structure of road networks, and thus further research is needed.

Chapter 3

Data

In this thesis, the data used was collected by previous research efforts (Zhu, 2010, Carrion and Levinson, 2012a). The main objective of this previous research is to understand the travel behavior of travelers due to the collapse of the I-35W bridge (August 1st 2007), after the replacement for the previously collapsed I-35W bridge opened to the public (September 18th 2008) in the Minneapolis-St. Paul region. The data consists of GPS observations, and web-based surveys collected before, and after the replacement bridge opened.

For this thesis, the chapters 4, 5, and 6 use this common data source, but the data sets (i.e. number of subjects and observations) are different as not all of them fulfill the requirements for each of these studies. The methodological requirements along with descriptive statistics for the whole data set, and for the data sets of the chapters 4, 5, and 6 are described subsequently.

3.1 Recruitment

Subjects were recruited through announcements posted in different media including: *Craigslist.org*, and *CityPages.com*; the free local weekly newspaper *City Pages*; flyers at grocery stores; flyers at city libraries, postcards handed out in downtown parking ramps; flyers placed in downtown parking ramps; and emails to more than 7000 University of Minnesota staff (students and faculty were excluded). More than 900 subjects responded, and consequently subjects were randomly selected among those that met the following requirements:

1. Age between 25-65,
2. Legal driver,
3. Full-time job and follow a “regular” work schedule
4. Travel by driving alone
5. Likelihood of being affected by the reopening of the new I-35W Mississippi River bridge.

Potential subjects (randomly selected from the respondents’ pool) were instrumented with GPS devices by two data collection efforts. The first was headed by Dr. Randall Guensler of the Georgia Institute of Technology and the subcontractor Vehicle Monitoring Technologies (VMTINC). Also, a local subcontractor (MachONE) was employed to instrument the subjects’ vehicles with GPS devices two weeks before the new I-35W bridge reopened. These GPS devices recorded the coordinates of the instrumented vehicle at every second between engine-on and engine-off events. The coordinates log collected by the GPS was transmitted to the server in real time through wireless communication. The subjects remained instrumented for 13 weeks without following any instructions with the exception of filling periodic surveys. The second was headed by the author and others affiliated with the University of Minnesota, Twin Cities Campus. The subjects were instrumented with logging-type GPS devices (QSTARZ BT-Q1000p GPS Travel Recorder powered by DC output from in-vehicle cigarette lighter) also approximately two weeks before the replacement I-35W bridge opened to the public. These GPS devices recorded the position of the instrumented vehicle at a frequency of 25 meters per location point registered between engine-on and engine-off events. These subjects remained instrumented for 8 weeks, during this time period the subjects followed their usual commute pattern without any instruction with the exception of filling periodic surveys. In addition, at the end of the study period (i.e. 8 weeks or 13 weeks depending on the data collection effort), subjects completed a comprehensive final web-based survey to evaluate the driving experience on routes using different bridges choices, provide socio-demographic information, and also answer some questions regarding route preferences.

A total of approximately 143 (46 by VMTINC, and 97 by University of Minnesota) subjects had usable (complete day-to-day GPS information) data required for this analysis.

3.2 GPS Data Processing

The raw data generated by the GPS devices consisted of a list of codes with detailed trip information including: record ID, latitude and longitude, date and time, and instantaneous speed. Each of the codes represent one point per 25 meters (or less) in the travel trajectories of each vehicle. In ideal conditions, the displacement of the vehicles is captured accurately by the GPS. In some situations, the records are not accurate, because it might take the GPS device a few minutes to initialize after the vehicle's engine is on. These points were excluded from the dataset. The actual routes used for the analysis were built by merging these points with a GIS map. This map is referred to as the *TLG network*, which is maintained by the Metropolitan Council and The Lawrence Group (TLG). It covers the entire 7-county Twin Cities Metropolitan Area and is the most accurate GIS map of this network to date. The TLG network contains 290,231 links, and provides an accurate depiction of the entire Twin Cities network at the street level. Twenty-meter buffers are used for all roads, in order clip the GPS records. All points outside of Twin Cities area as well as off-road points were excluded. The remaining points were regrouped into trips; these trips contained all points between one engine-on and engine-off events for each subject. In this way, all trips by each subject were identified along with the characteristics of each trip, including the starting time, the ending time, the path used, and travel speed on each link segment along the route. This complete process was done inside the ArcGIS environment.

3.3 Surveys

Web-based surveys are used for collecting profiles, attitudes, and stated preferences (SP) of the subjects. These offer significant advantages over paper-based surveys:

- reduced computational time spent processing the data;
- use of audiovisual features;
- restrictive control of answers (e.g. leaving questions blank);
- less active participation of experimenters; and others.

For this research, three types of web-based surveys were employed. The first survey filtered the prospective participants for the experiment according to the requirements listed in Section 3.1. The second survey captured subject's weekly perceptions of bridge attributes (e.g. congestion level) for morning and afternoon commutes. The third survey collected the final stated preferences at the end of the study period (i.e. 8 weeks or 13 weeks depending on the data collection effort). This survey included questions about: socio-demographics (e.g. age, income); perceived attributes (e.g. travel time predictability) for both morning and afternoon commutes; bridge preferences for morning and afternoon commutes; reasons (e.g. travel time) for selecting a route instead of others; stating threshold of willingness to pay a toll cost (using only HOT lanes) for distinct travel time savings; and stating threshold of willingness to pay for distinct travel time reliability savings.

The weekly web-based survey was completed by the study participants each Monday, Wednesday and Friday. In contrast the final survey was completed only at the end of the experiment. Examples of these surveys are in the appendix. Also, only the 97 subjects of the University of Minnesota reported the travel times of their trips in periodic surveys. This periodic survey is included in the appendix.

3.4 General Methodology

The data analysis process required for chapters 4, 5, and 6 is described subsequently.

1. Identification of commute trips per subject from GPS data on the bridges of interest (see figure 3.1);
2. Information extraction (e.g. travel time) of commute trips per subject from GPS and survey data;

3. Specification and estimation of econometric models using the extracted information from GPS and from survey data.

The first phase uses the coordinates (latitude and longitude) of the trips per subject, and the *TLG network* in order to identify the trips crossing bridges, and the bridges crossed. Only commute trips crossing bridges are considered in this thesis. The identification is done by spatial matching the coordinates of each bridge of interest (see figure 3.1) to the coordinates of each set of trips for each subject. Also, subjects' trips must start at their home (work) and end at their work (home) locations in order to be considered commute trips (only *direct* commute trips). The distance tolerance between origins (destinations) to home (work) locations was set to 600 meters. In addition, a threshold was set for the start of a new trip at 5 minutes. This temporal constraint guarantees that the trips are mostly direct, and avoids confounding difficulties such as chained trips. The home and work locations are geocoded (transformed into latitude and longitude coordinates) from the actual addresses provided by the subjects on the web-based surveys. The origin and destination pair of each trip is obtained by mapping the coordinate points into trajectories of engine-on and engine-off events. Moreover, inaccurate points due to GPS "noise", and out-of-town trips (e.g. during Thanksgiving) were excluded. Lastly, only the trips after September 18th are considered as this is the date the new I-35W Bridge opened to the public at 5 AM. Thus, all the bridges of figure 3.1 are available to the subjects.

For chapters 4 and 5, it is required to match the commute trips from GPS data to the commute trips from survey data. This is required to generate the dependent variable as it is defined in the section 4.2.1 of chapter 4. Also, this allows to obtain day-to-day travel time distributions as described in the section 5.2 of chapter 5. These day-to-day travel time distributions are obtained for the *most used* bridges of freeways, and *most used* bridges of arterials for each of the subjects. This is accomplished by matching the dates of commute trips from GPS data to commute trips from survey data. The subjects completed the information of commute trips in the periodic surveys within the same day that they took their trips. Thus, times of departure of the commute trips must be earlier than the time of completing the periodic survey by the subjects. Furthermore, any trips that are not considered commute trips according to the subjects in the periodic survey data are excluded. The only observations considered are those of subjects with self reported travel times of their trips in periodic surveys. Only 64 subjects of the University of Minnesota sample has the required data for chapter 4, and only 39 subjects of the University of Minnesota sample for chapter 5.

Chapter 6 does not use information from periodic surveys. Only uses GPS data from subjects, and socio-demographics from the final survey. In addition, only morning trips (those between 4 AM and 11 AM) are considered, because it is suspected that subjects are not able to gather information from other non-commute trips during the morning, especially when the subjects drive directly from home to work without any side stops. This is further elaborated in section 6.3 of chapter 6. Only 65 subjects (26 from VMTINC, and 39 from University of Minnesota) has the required data for this chapter.

The second phase extracts usable information (e.g. travel times, gender) from the matched trips. This process is performed for both home to work trips, and work to home trips.

The third phase consists of using the extracted information as regressors and dependent variables in each of the econometric models: linear regression models, and logistic regression models (chapter 4); random utility models (chapter 5); and duration models (chapter 6).

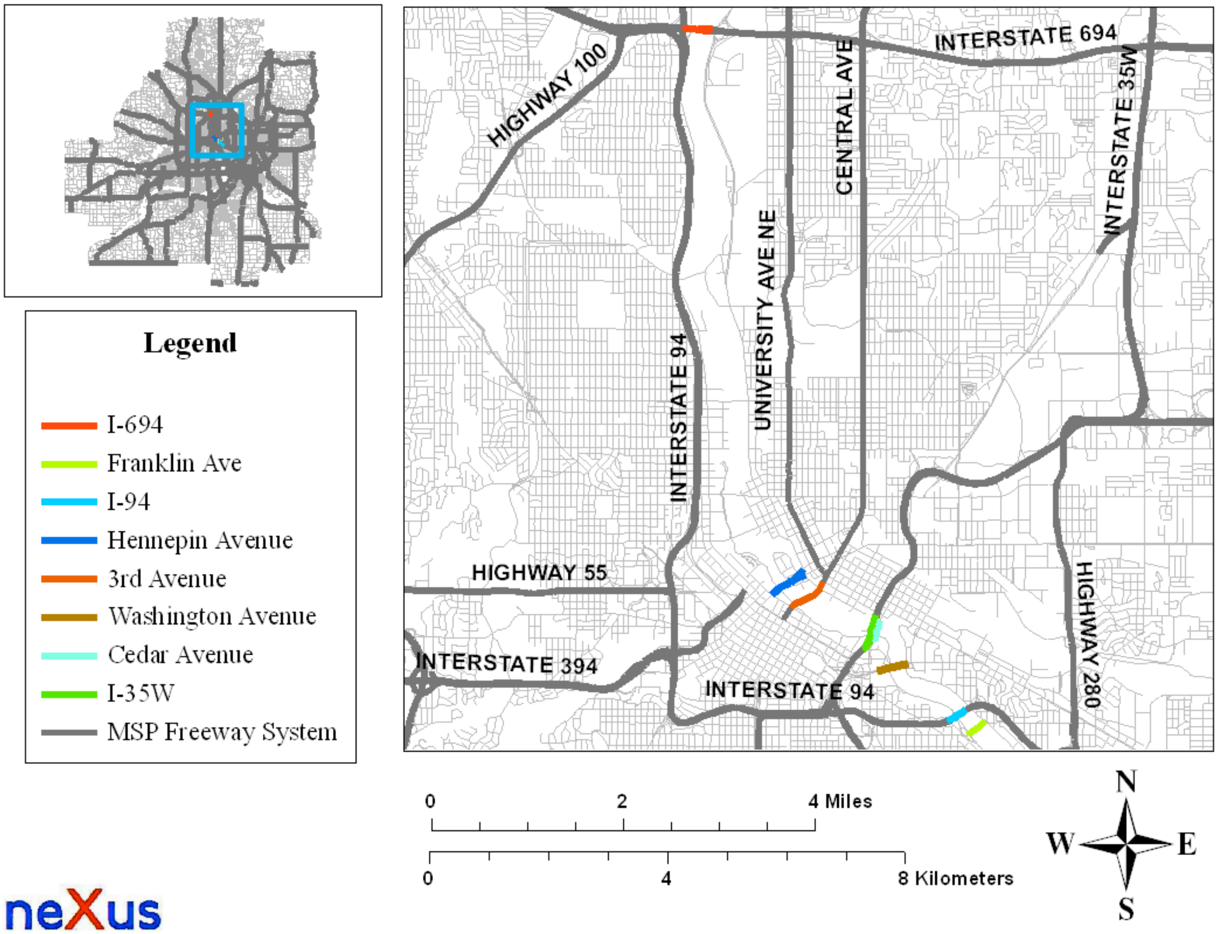


Figure 3.1: Bridge locations (Source: Carrion and Levinson (2012a))

3.5 Descriptive statistics of the data sources

3.5.1 Socio-Demographics

Table 3.1 summarizes socio-demographic information of the subjects in the complete sample: the subjects in the VMTINC sample, and the subjects in the University of Minnesota sample (UMN). The sample differs from the population of the Minneapolis-St. Paul region in several ways: more female subjects, subjects are older, more educated, and have a more uniform distribution of income. The main difference between both samples is that subjects of the VMTINC sample are more educated, and have a higher female proportion than subjects of the University of Minnesota sample.

Table 3.1: Socio-Demographics attributes of data sources

Number of Subjects		143	46	97	
		Complete	VMTINC	UMN	Twin Cities
Sex	Male	41.25%	36.96%	43.29%	49.40%
	Female	58.75%	63.04%	56.71%	50.60%
Age (Mean, Std. Deviation)		(47.86, 8.86)	(49.38, 10.42)	(46.34, 7.29)	(34.47, 20.9)
Education	11th grade or less	0.00%	0.00%	0.00%	9.40%
	High School	13.09%	4.35%	17.24%	49.60%
	Associate	24.99%	30.43%	22.41%	7.70%
	Bachelors	45.22%	36.96%	49.14%	23.20%
	Graduate or Professional	16.69%	28.26%	11.21%	10.10%
Household Income	$49,999 or less	20.20%	28.26%	16.38%	45.20%
	$50,000 to $74,999	30.73%	28.26%	31.90%	23.30%
	$75,000 to $99,999	29.44%	26.09%	31.03%	14.60%
	$100,000 to $149,999	17.06%	13.04%	18.97%	11.00%
	$150,000 or more	4.17%	4.35%	2.59%	3.16%
Race	Black/African American	7.36%	6.52%	7.76%	6.20%
	White or Caucasian	83.06%	78.26%	85.34%	87.70%
	Others	9.58%	15.22%	6.90%	6.10%

Minneapolis' Population statistics are obtained from the US Census Bureau, (cen)

3.6 Descriptive statistics of the data sets

This section presents descriptive statistical analyses of the data sets of chapters 4, 5, and 6. These analyses are: socio-demographic summary statistics for the data sets of each chapter; bar graphs of travel times for the data set of chapter 4; kernel density estimates of travel time distributions for the data set of chapter 5; and nonparametric estimates of the survivor function and the cumulative hazard function for the data of chapter 6.

3.6.1 Socio-Demographics

Table 3.2 summarizes socio-demographic information of the subjects in the data sets for each chapter. The data sets of chapters 4, and 5 differ from the population of the Minneapolis-St. Paul region in several ways: subjects are older, more educated and have a more uniform distribution of income. The data set of chapter 6 is similar to the others, except that the proportion of females is significantly greater than the proportion of males.

Table 3.2: Socio-Demographics attributes of the data sets

Number of Subjects		64	39	65	
		Chapter 4	Chapter 5	Chapter 6	Twin Cities
Sex	Male	43.93%	46.15%	27.69%	49.40%
	Female	56.07%	53.85%	72.31%	50.60%
Age (Mean, Std. Deviation)		(52.38, 9.91)	(50.44, 10.81)	(51.11, 10.31)	(34.47, 20.9)
Education	11th grade or less	0.00%	0.00%	0.00%	9.40%
	High School	18.03%	17.95%	15.87%	49.60%
	Associate	22.54%	15.38%	23.81%	7.70%
	Bachelors	51.56%	61.54%	46.03%	23.20%
	Graduate or Professional	7.86%	5.12%	14.29%	10.10%
Household Income	$49,999 or less	15.26%	23.08%	26.15%	45.20%
	$50,000 to $74,999	23.47%	20.51%	20.00%	23.30%
	$75,000 to $99,999	36.53%	25.64%	32.31%	14.60%
	$100,000 to $149,999	21.16%	23.07%	18.46%	11.00%
	$150,000 or more	3.58%	7.69%	3.08%	5.90%
Race	Black/African American	11.33%	7.69%	9.52%	6.20%
	White or Caucasian	81.73%	79.49%	76.19%	87.70%
	Others	6.93%	12.81%	14.28%	6.10%

Minneapolis' Population statistics are obtained from the US Census Bureau, (cen)

3.6.2 Statistics of travel time

Figures 3.3, and 3.2 summarize trips according to travel times reported (stated, and expected) by the subjects in the periodic surveys, and measured through GPS devices on the subjects' vehicles. These travel times from the surveys and GPS devices are solely part of the data set of chapter 4.

Figure 3.2 compares proportions of trips in the following order: GPS travel times vs. Stated travel times; Expected travel times vs. GPS travel times; and Stated travel times vs. Expected travel times. In general, subjects' stated travel times are greater than measured travel times for trips with travel times less than 20 minutes. In contrast, subjects' stated travel times are smaller than measured travel times for trips with travel times more than 25 minutes. Furthermore, subjects' expected travel time are closely similar to measured travel times for trips with travel times less than 15 minutes. However, subjects expect higher travel times for trips with travel times between 15 minutes and 30 minutes. Lastly, trips with travel times greater than 40 minutes are always underestimated (both stated and expected) by the subjects. This findings agree with Vierordt (1868)'s law.

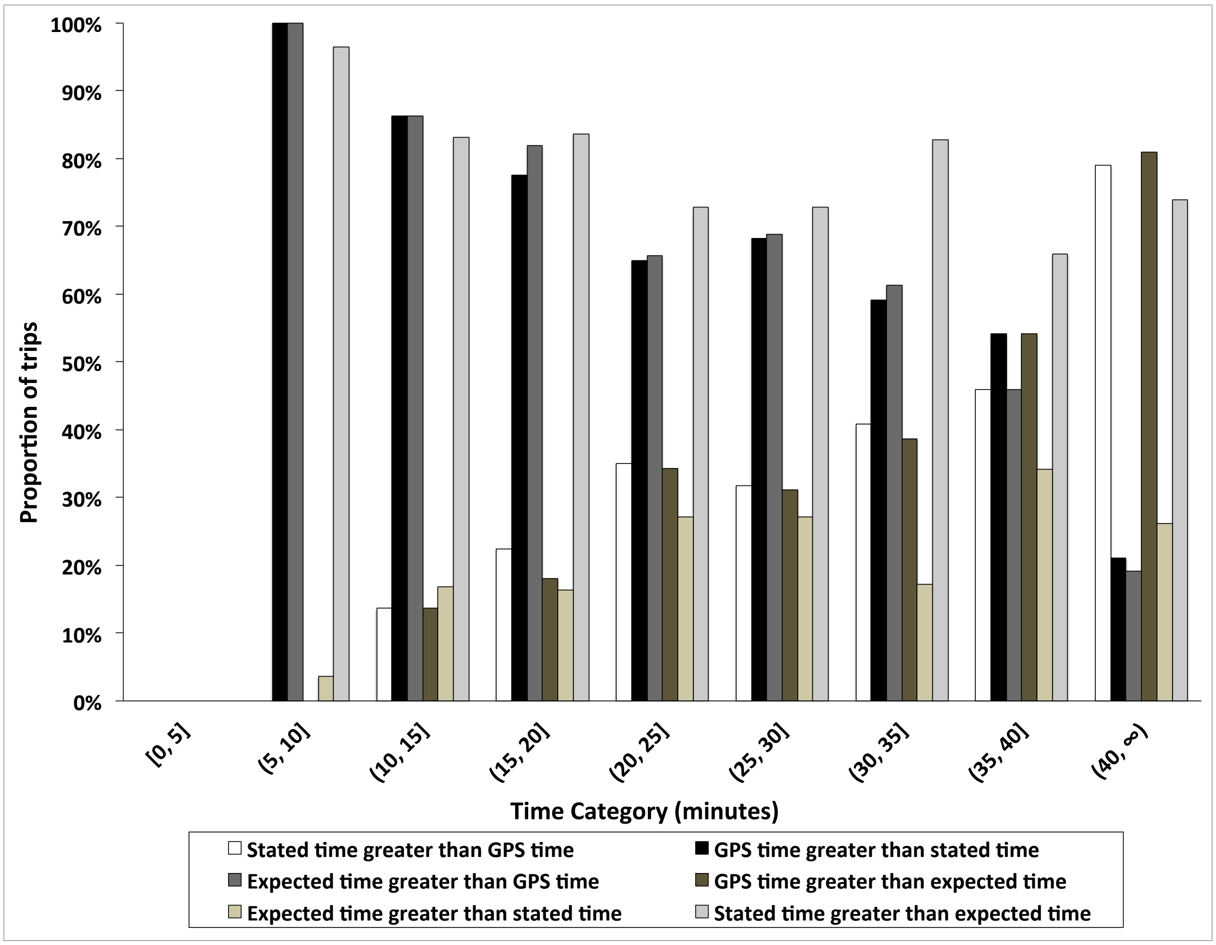


Figure 3.2: Proportion of trips according to travel time of commute from GPS data and survey data - GPS vs. Stated; Expected vs. GPS; Stated vs. Expected

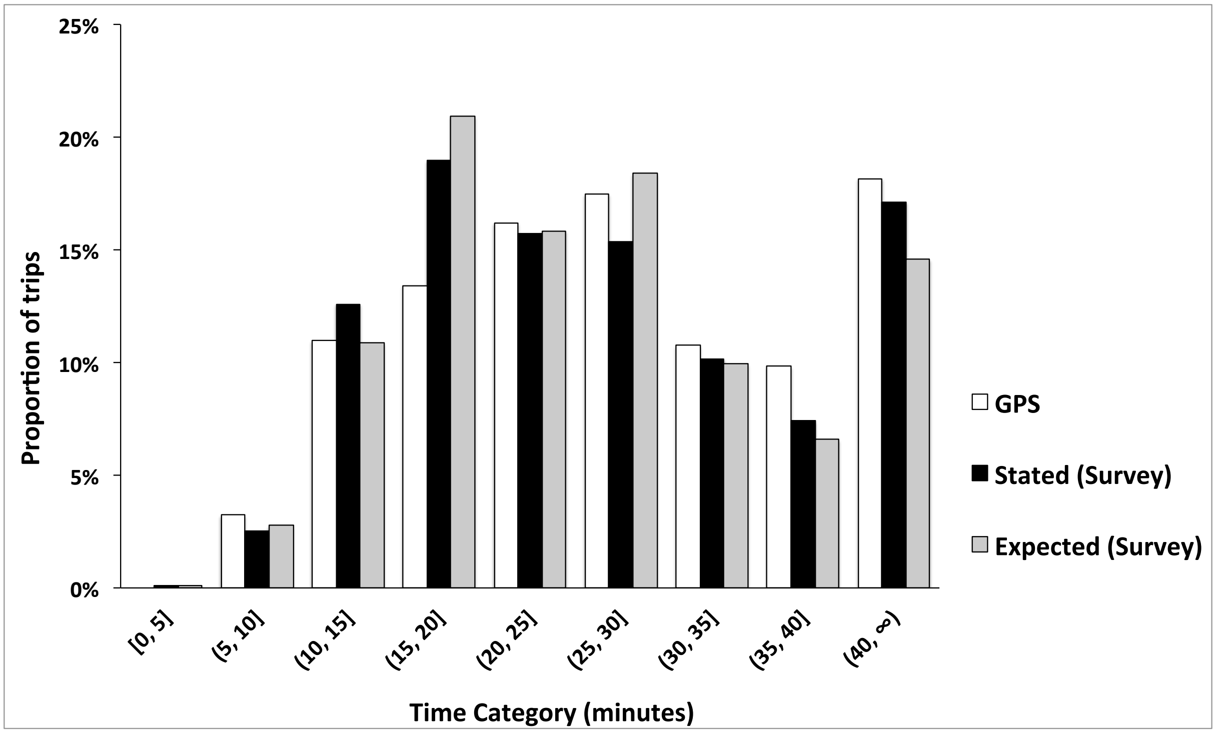


Figure 3.3: Proportion of trips according to travel time of commute from GPS data and survey data

3.6.3 Statistics of (mean) travel time distributions

Figures 3.4, and 3.5 present kernel density estimates (Gaussian, 4.2 bandwidth) of (mean) travel time distributions of *most used* freeway bridges, and *most used* arterial bridges by the subjects. These travel times of freeway bridges and arterial bridges are solely part of the data set of chapter 5.

The term *most used* (see section 5.2 in chapter 5) refers to the bridge with the highest number of commute trips. The term *mean travel time distribution of most used freeway (arterial) bridge* refers to a distribution obtained by aggregating each mean computed on each day to day travel time distribution of the *most used* freeway (arterial) bridge for each subject. The figures indicate that the mean subjective travel time distribution (i.e. travel times as reported by a traveler in the periodic surveys; these travel times are stated, and are not expected), does not matches the mean objective travel time distribution (i.e. travel times as measured by GPS devices). It has different variability, and centrality. Thus, perception error is indeed present as both should be significantly similar if there were no perception error.

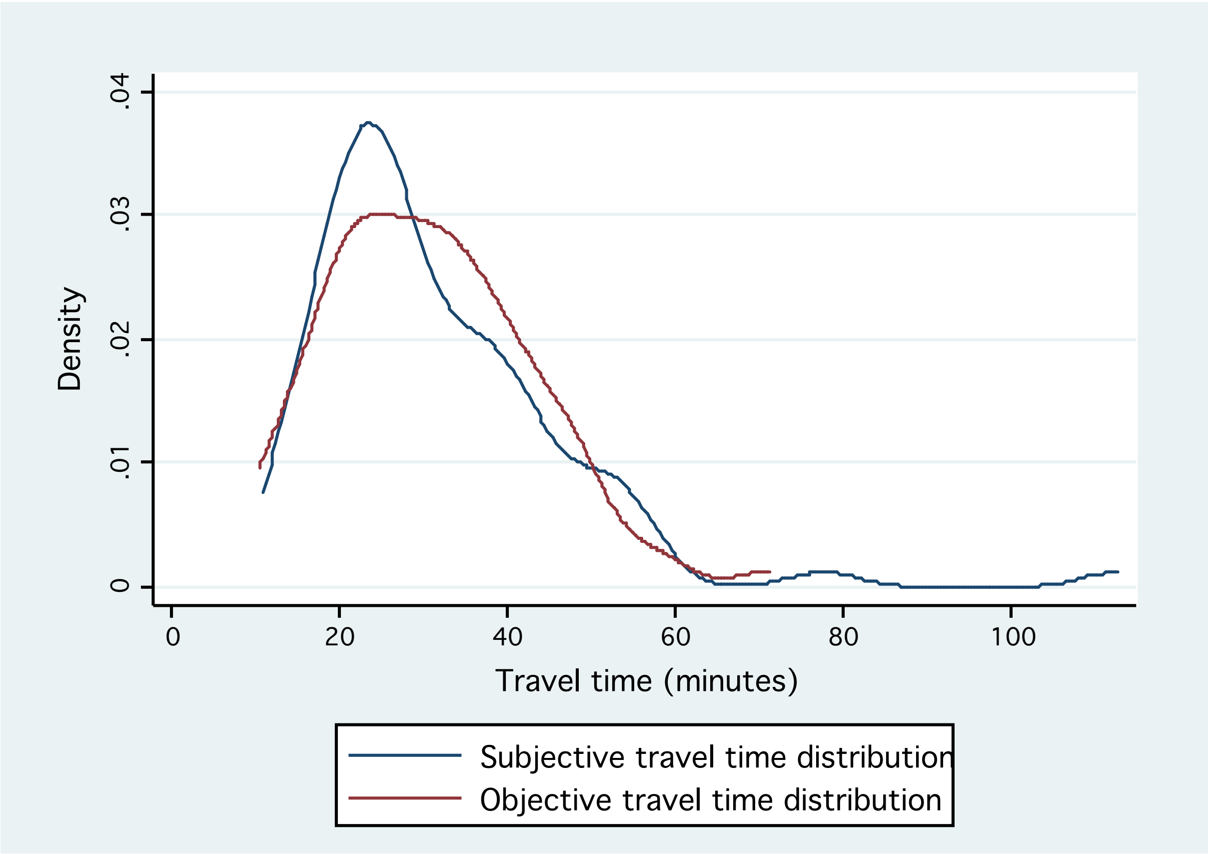


Figure 3.4: Kernel density estimates (Gaussian, 4.2 bandwidth) of (mean) travel time distributions of most used freeway bridges by the subjects

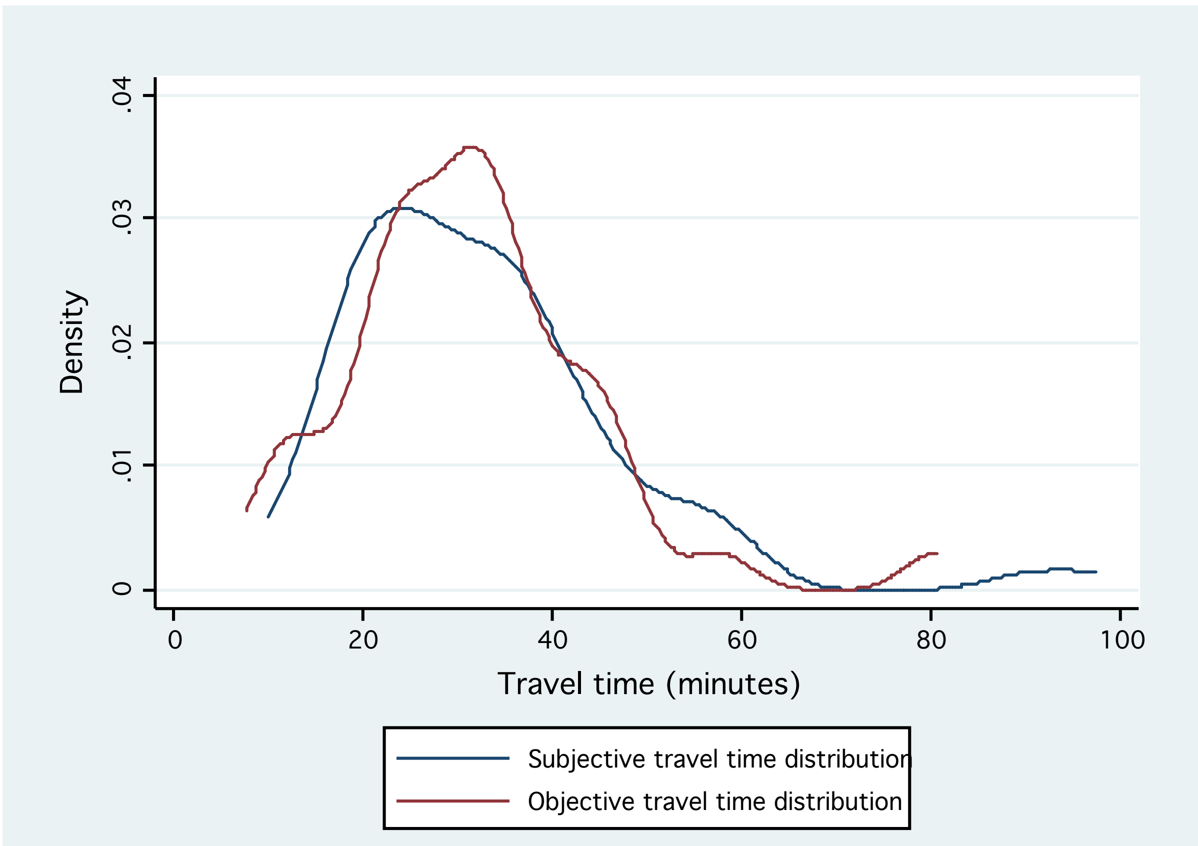


Figure 3.5: Kernel density estimates (Gaussian, 4.2 bandwidth) of (mean) travel time distributions of most used arterial bridges by the subjects

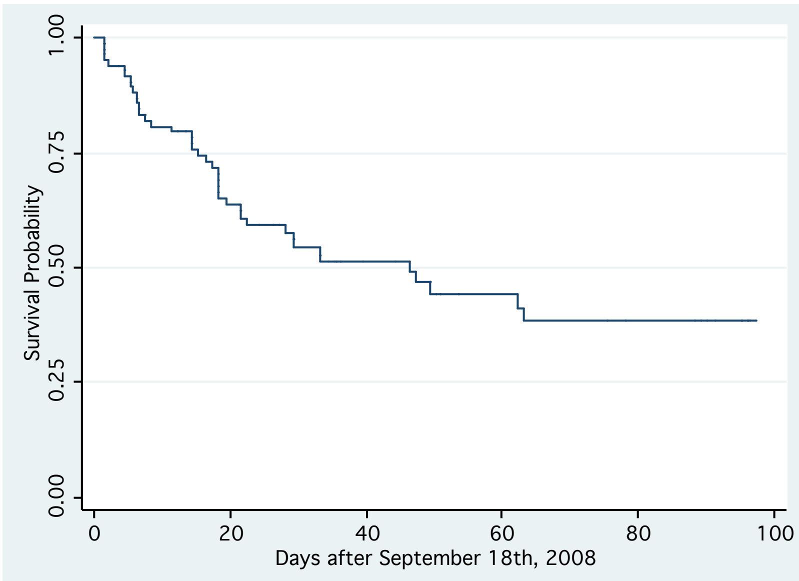
3.6.4 Nonparametric analysis

Figures 3.6a and 3.6b present nonparametric estimates of the survivor function (i.e. Kaplan Meier), and the cumulative hazard function (i.e. Nelson-Aalen) of the subjects' *single-spell durations*. These durations are solely part of the data set of chapter 6.

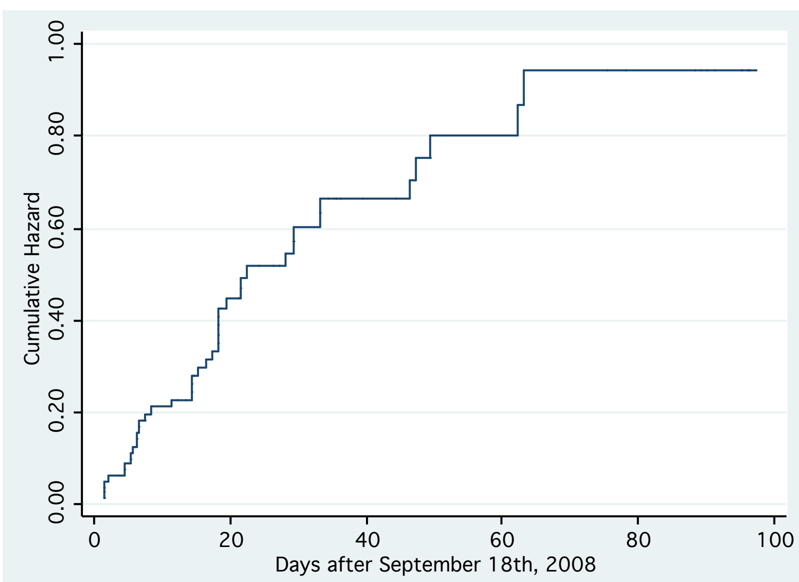
A *single-spell duration* is defined as the date and time elapsed from September 18th 5:00 AM (I-35W bride reopened to the public) until the date and time a subject *consistently* leaves his *current bridge choice* for any other of those in figure 3.1 (transition is observed), or the date and time until the GPS device is retrieved from the subject, and the subject has not left his *current bridge choice* (transition is not observed). The term *current bridge choice* refers to the subject's bridge choice at or after September 18th 5:00 AM. The term *consistently* refers to a subject's transition from his current bridge choice to another bridge at least two times consecutively. The term *single-spell duration* refers to modeling only one single transition from the *current bridge choice* to any other of those in figure 3.1. This single transition is only the first transition observed in the subjects' GPS data. Readers should consult section 6.3 of chapter 6. The figures present the shape of the survival curve, and the cumulative hazard curve, which are estimated based on the data of *single-spell durations* of subjects without considering any covariates (see section 6.3.3 of chapter 6). The shape is obtained through a product limit estimator (Kalbfleisch and Prentice, 2002, Cameron and Trivedi, 2005) similar to the one discussed in section 6.3.5 of chapter 6. The Nelson-Aalen cumulative hazard (figure 3.6b) indicates that there is rapid growth in the susceptibility to leave the *current bridge choice*, and that eventually it flattens. This rapid growth happens before 40 days after the date September 18th 2008. This agrees with the Kaplan-Meier survival curve (figure 3.6a) that indicates a sharp drop in the survival probability by 20 days after September 18th 2008, and a smoother drop between 20 days after September 18th 2008, and 40 days after September 18th 2008.

Figure 3.6: Nonparametric curves

(a) Kaplan-Meier survival curve estimate of subjects' single-spell duration



(b) Nelson-Aalen cumulative hazard estimate of subjects' single-spell duration



Chapter 4

Empirical functions of travel time perception

4.1 Introduction

In the transportation research literature, subjects' perception of travel times has been found to be a significant factor in studies (Levinson et al., 2004, 2006, Peer et al., 2010, Peer, 2013). Travelers overestimate or underestimate the actual travel times of their trips. Recent research (Parthasarathi, 2011, Parthasarathi et al., 2012) is just starting to systematically "unpack" the factors governing the *perception error* of the travelers. In this study, the primary objective is to further uncover the factors governing the *perception error* along with the nonlinearities (e.g. functional forms) linking the *perception error*, and the factors. The methodology is based on regression analysis on data collected (surveys, and Global Positioning System [GPS] points) of commuters recruited from a previous research study in the Minneapolis-St. Paul region (Zhu, 2010, Carrion and Levinson, 2012a). The main characteristics of the data are that actual route information as in spatial location of subjects in the network is known, home and work location of subjects are known, and also subjects' information regarding their travel experience, and their time restrictions (i.e. travelers are allowed to arrive late to work without any reprehension). Furthermore, only direct (i.e. no trip chaining) commute trips (from home to work, and from work to home) are considered for the analysis. For these work trips, the subjects' self-reported travel times, and the subjects' travel times measured by GPS devices were collected. Also, the commute trips are trips crossing the Mississippi River using any of the bridges in figure 3.1.

4.2 Econometric models

In this study, the dataset is analyzed through *regression analysis* (Johnston and DiNardo, 1997, Cameron and Trivedi, 2005, Wooldridge, 2010, Greene, 2012). The dataset is composed of several observations (i.e. home to work trips, and work to home trips) per subject. There are 64 distinct subjects (see section 3.6), and on average 13.5 observations per subject. There is a total of 865 observations (see section 3.1). These observations correspond to the set of home to work trips per subject (514 trips), and the set of work to home trips per subject for the same trips (351 trips; see section 3.4) from GPS and survey data per subject. In addition, these work trips (or observations) are trips crossing the Mississippi River using any of the bridges in figure 3.1.

4.2.1 Linear regression models

The dependent variable (τ_{jn}) is defined as the ratio of reported travel time from surveys (T_{rjn}) and measured travel time from GPS data (T_{mjn}). This ratio is calculated for every observation n in the set of observations $\mathcal{J}_n$ for subject n, and for every subject n in the set of subjects $\mathcal{N}$. This is the exact definition used in previous research (Parthasarathi, 2011, Parthasarathi et al., 2012). This ratio is defined mathematically as,

$$\tau_{jn} = \frac{T_{rjn}}{T_{mjn}} \tag{4.1}$$

The observations (i.e. trips) are divided into two groups: overestimated trips; and underestimated trips. The former refers to trips by subjects with $\tau_{jn} > 1$. The latter refers to trips by subjects with $\tau_{jn} < 1$. The linear regression models described subsequently are estimated on each of these two groups.

The general structure of the linear regression models follows the *random effects model* for panel data (Johnston and DiNardo, 1997, Cameron and Trivedi, 2005, Wooldridge, 2010, Greene, 2012). This structure is used to handle the correlations due to unobservable variables across observations belonging to the same subject. The α_n term captures the correlations (i.e. $Cov(\tau_{jn}\tau_{jn'}) = E(\alpha_n\alpha_n) = \sigma_\alpha^2$) across observations ($j \in \mathcal{J}_n$) of the same subject n. Mathematically, the general structure is

$$\tau_{jn} = f(\mathbf{x}_{jn}, \mathbf{z}_{jn}; \beta) + \alpha_n + \epsilon_{jn} \tag{4.2}$$

where

- $f(\cdot)$: Functional forms for the covariates.
- τ_{jn}: The dependent variable defined in equation 4.1.
- $\mathbf{x}_{jn}$ The vector of continuous covariates described in section 4.2.3.
- $\mathbf{z}_{jn}$ The vector of categorical covariates described in section 4.2.3.
- β: The vector of parameters to be estimated.
- $\alpha_n \sim$ i.i.d. $N(0, \sigma_\alpha^2)$ for all n
- $\epsilon_{jn} \sim$ i.i.d. $N(0, \sigma_\epsilon^2)$ for all j and n

In this study, the functional forms (i.e. $\tau_{jn} = f(\mathbf{x}_{jn}; \beta)$) used for the covariates are: linear, quadratic, and Cobb-Douglas. Mathematically, these functional forms are:

Linear: $\tau_{jn} = \sum_{i=0}^{h} \beta_i x_{jn} + \sum_{i=h+1}^{k} \beta_i z_{jn}$

Quadratic: $\tau_{jn} = \sum_{i=0}^{h} \beta_i x_{jn} + \sum_{i=h+1}^{t} \beta_i (x_{jn})^2 + \sum_{i=t+1}^{k} \beta_i z_{jn}$

Cobb-Douglas: $\tau_{jn} = \beta_0 \prod_{i=1}^{h} (x_{jn})^{\beta_i} e^{\sum_{i=h+1}^{k} \beta_i z_{jn}}$

where

- τ_{jn}: The dependent variable defined in equation 4.1.
- x_{jn} The vector of continuous covariates described in section 4.2.3.
- z_{jn}: The vector of categorical covariates described in section 4.2.3.
- β: The vector of parameters to be estimated.

It should be noted that k is the total number of covariates, and that $h + t = k$.

4.2.2 Logistic regression model

The dependent variable (δ_{jn}) is defined as $\delta_{jn} = 1$ if $\tau_{jn} > 1$, and $\delta_{jn} = 0$ if $\tau_{jn} < 1$. $\tau_{jn} > 1$ represents the trips (i.e. observations) of subjects that *overestimated* their travel times. In contrast, $\tau_{jn} < 1$ represents the trips (i.e. observations) of subjects that *underestimated* their travel times. Mathematically, the dependent variable is

$$\delta_{jn} = \begin{cases} 1, & \text{if } \tau_{jn} > 1 \\ 0, & \text{if } \tau_{jn} < 1 \end{cases} \tag{4.3}$$

The general structure of the logistic regression models follows the *random effects model* for panel data (Johnston and DiNardo, 1997, Cameron and Trivedi, 2005, Wooldridge, 2010, Greene, 2012). This structure is used to handle the correlations due to unobservable variables across observations belonging to the same subject. Similarly to the linear regression models, the α_n term captures the correlations across observations ($j \in \mathcal{J}_n$) of the same subject n. Mathematically, the general structure is

$$Prob[\delta_{jn} = 1 | \mathbf{x}_{jn}, \mathbf{z}_{jn}, \alpha_j] = \frac{e^{f(\mathbf{x}_{jn}, \mathbf{z}_{jn}, \alpha_n; \beta)}}{1 + e^{f(\mathbf{x}_{jn}, \mathbf{z}_{jn}, \alpha_n; \beta)}} \tag{4.4}$$

where

- $f(\cdot)$: Functional forms for the covariates.

- δ_{jn}: The dependent variable defined in equation 4.3.
- $\mathbf{x}_{jn}$ The vector of continuous covariates described in section 4.2.3.
- $\mathbf{z}_{jn}$ The vector of categorical covariates described in section 4.2.3.
- β: The vector of parameters to be estimated.
- $\alpha_n \sim$ i.i.d. $N(0, \sigma^2)$ for all n

4.2.3 Covariates

The covariates are obtained from the periodic web-based surveys, and/or from the GPS data. Readers may refer to section 3.4 for details. Each of the covariates are defined subsequently, and also the covariates are related to time perception research from psychology in the subsequent section.

Arrival Flexibility:

It is a set of four binary variables. It represents the subjects workplace arrival time constraints. The categories are: had to be there at the work start time; may arrive within 20 minutes of the work start time; may arrive within 60 minutes of the work start time; may arrive at any time past the work start time.The first category is the base case. It is obtained from the periodic surveys.

Type of trip:

It is a binary variable indicating whether the trip originates from home (1 = from home to work) or from work (0 = from work to home).

Expected travel time of the trip:

It is the travel time a subject indicates as *expected* to arrive at their destination. This is different from the reported travel times as those are based on the a subject's estimate of the actual travel time of the trip. Furthermore, this variable has the potential for being endogenous. Therefore, the proposed models (see section 4.2) are estimated twice: models including this variable; and models excluding this variable. It is in minutes. It is obtained from the periodic surveys.

Traffic information:

It is a binary variable indicating whether a subject received any type of pre-trip travel information. 1 = received information; 0 = did not receive information. It is obtained from the periodic surveys.

Trips on Interstate bridges:

It is a binary variable indicating whether a subject crossed the Mississippi River using any of the Interstate bridges in figure 3.1. 1 = used an Interstate bridge; 0 = did not use an Interstate bridge. It is obtained from the GPS data.

Relative discontinuity:

It is the sum of changes in street hierarchy (i.e. discontinuity) divided by the trip length. A change in street hierarchy is defined by the change of segment speed as it is illustrated in figure 4.1. For example, a traveler moves from origin A to destination B. The traveler moves from a link with hierarchy k_1 to a link with hierarchy k_2, and so on. It is calculated on the actual commute routes of each trip taken by each subject from the GPS data (Xie and Levinson, 2007, Parthasarathi, 2011).

Mathematically, a change in street hierarchy a from link with hierarchy k_1 to a link with hierarchy k_2 is

$$y_a = |k_1 - k_2| \tag{4.5}$$

The discontinuity of the trip along a path P is

$$Y_P = \sum_{a \in P} y_a \tag{4.6}$$

The relative discontinuity is

$$Y_p^{'} = \frac{Y_P}{l_P} \tag{4.7}$$

where

- Y_P: Discontinuity of the trip along a path P
- l_P: Length (km) of trip along a path P

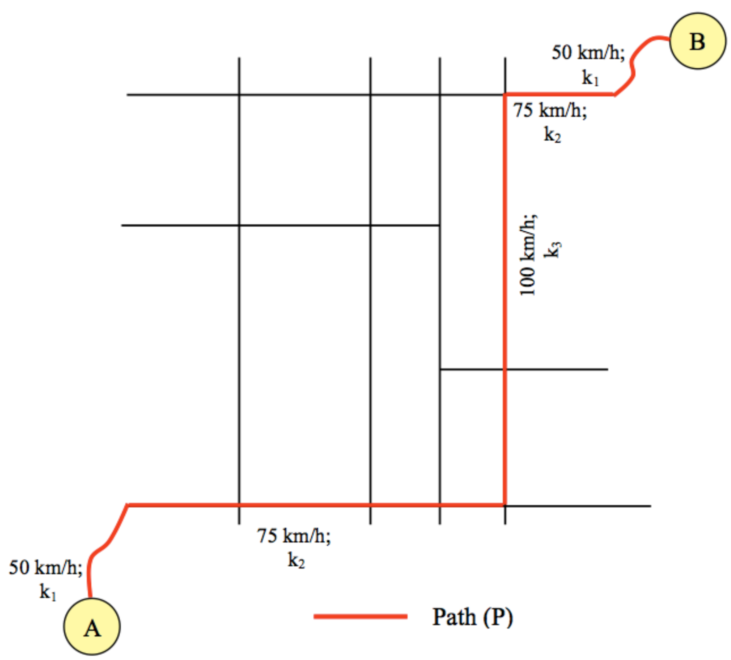


Figure 4.1: Illustration of trip discontinuity (Source: Parthasarathi (2011))

Proportion of limited access roads:

This measure represents the presence of limited access roads. These roads are devoid of traffic signals, and may be grade separated, and thus are dissimilar from other roads.

It is obtained by diving the trip length of the trip on limited access roads to the total length of the trip. It is calculated on the actual commute routes of each trip taken by each subject from the GPS data (Xie and Levinson, 2007, Parthasarathi, 2011).

Mathematically, it is defined as

$$H_P = \frac{l_{la}}{l_P} \tag{4.8}$$

where

- H_P: Proportion of limited access roads of path P
- l_{la}: Length (km) of trip on limited access roads along a path P
- l_P: Length (km) of trip along a path P

Proportion of signalized arterials:

This measure represents the presence of signalized arterials. Typically, these roads have traffic signals and/or traffic signs, and may be used to access several types of destinations (e.g. commercial).

It is obtained by diving the trip length of the trip on signalized arterials to the total length of the trip. It is calculated on the actual commute routes of each trip taken by each subject from the GPS data (Xie and Levinson, 2007, Parthasarathi, 2011).

Mathematically, it is defined as

$$A_P = \frac{l_{sa}}{l_P} \tag{4.9}$$

where

- A_P: Proportion of signalized arterials of path P
- l_{sa}: Length (km) of trip on signalized arterials along a path P
- l_P: Length (km) of trip along a path P

Circuity:

It is defined as the ratio of the network distance of a path P to the Euclidean distance of the origin and destination corresponding to the path P. This is illustrated in figure 4.2. This measure captures the inefficiency in the network from a traveler's perspective. It is calculated on the actual commute routes of each trip taken by each subject from the GPS data (Xie and Levinson, 2007, Parthasarathi, 2011). It is unitless.

Mathematically, it is defined as

$$C_P = \frac{l_d}{l_e} \tag{4.10}$$

where

- C_P: Circuity of path P
- l_d: Network distance (km; trip length) of trip along a path P
- l_e: Euclidean distance (km) of trip along a path P

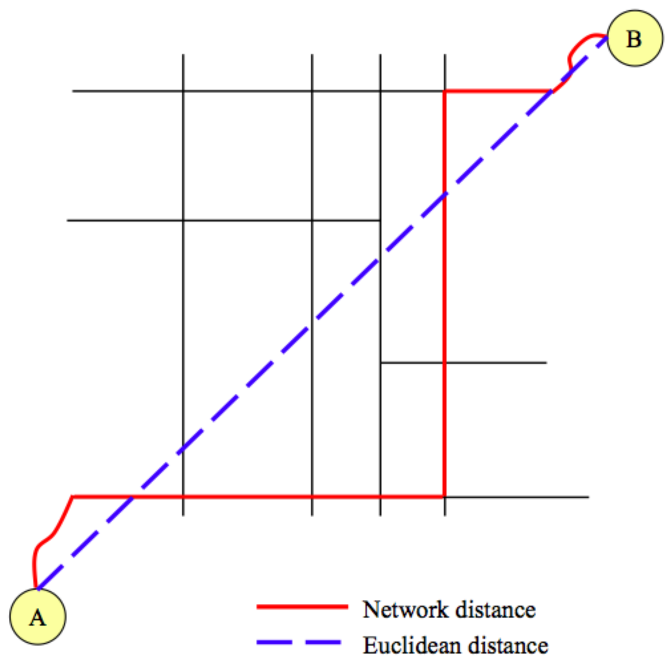


Figure 4.2: Illustration of trip circuity (Source: Parthasarathi (2011))

Congestion level:

It is set of three binary variables. It represents the subjects description of their travel experience with regards to the experienced congestion during their trips. The categories are: low congestion (i.e. not at all congested or very little congestion); medium congestion (i.e. average congestion level); high congestion (i.e. very congested). The first category is the base case. It is obtained from the periodic surveys.

Stress level:

It is set of three binary variables. It represents the subjects description of their travel experience with regards to their stress level during their trips. The categories are: low stress (i.e. not at all stressful or lightly stressful); medium stress (i.e. average stress level); high stress (i.e. very stressful). The first category is the base case. It is obtained from the periodic surveys.

Fear of driving on the I-35W bridge and other bridges in the vicinity:

It is a binary variable. This variable identifies the subjects that admitted they avoid bridges (including the I-35W bridge, Washington Ave bridge, and 10th Street bridge), because of fear of bridge collapse as indicated in the periodic web-based surveys.

Gender

It is a binary variable; 1 = Male; 0 = Female. It is obtained from the periodic surveys.

Income

It is a set of three binary variables: Low income ($[\$0, \$49,999]$), Medium income ($[\$50,000, \$99,999]$). and High income ($[\$100,000, \infty+)]$). The first category is the base case (2008 US dollars).

It is obtained from the periodic surveys.

4.2.4 Hypotheses of the study

The following hypotheses are based on studying time perception research from psychology. This research is discussed in section 1.1 of chapter 1. In summary, the research of time perception from psychology has classified the perception of time into three main categories: subjective time passage (i.e. perception of the speed that time passes); estimation of time duration; and simultaneity and succession of time. The estimation of time duration is the most frequently studied category by psychologists, and thus it is better understood. It is also the dimension of time perception that will be focus of this study, and it has been the focus of most studies investigating perception of travel time in the transportation literature. Moreover, psychologists have identified several factors influencing the time duration: temporal relevance, temporal uncertainty, affective elements, arousal, task complexity, temporal expectancies, absorption and attentional deployment. However, it must be emphasized that the variability of methods used by psychologists, and the inclusion of previously mentioned factors together makes it difficult to have a clear, and systematic understanding of the influence of these factors on the perception of time duration. Thus, psychologists are not entirely in agreement with the factors that influence the time duration. In addition, psychologists tend to conduct their studies in laboratory settings, and thus there is more control on the factors studied in comparison to real life situations.

In this study, the time perception research informs the expected direction of the covariates in the models. It is likely that each of these categories are related to attributes of the travelers' environment, and to attributes of the travelers themselves. During their spatial movement, it is plausible that travelers are undergoing different levels of effort, complexity, and other factors found to alter the perception of the duration of time. Thus, it is *conjectured* that the factors may be represented as more consolidated in a transportation environment. Furthermore, this study is not ideal for this research, because the data source is not specifically designed to address any research questions with regards to the perception of travel time. Thus, variables collected from the data are difficult to group specifically into any of the main factors in psychological research, and will just fit a less detailed set of categories. These less detailed set of categories are the *conjectured* categories discussed as follows:

1. **Temporal relevance**: It refers to the significance of time for performing a task in an optimal way. In other words, the level of importance and relevancy of time for a given task. The time perception research (Zakay, 1992) of psychology indicates that subjects will overestimate (underestimate) the time duration, because they pay more (less) attention to time when temporal relevance is high (low).

   In transportation, commuters with significant arrival flexibility will tend to underestimate their travel time, because they are likely to exhibit low *temporal relevance*. They are allowed to "waste time" during their trip to work. In addition, commuters that are driving from home to work will tend to overestimate their travel time, because they are likely to exhibit high *temporal relevance*. They are bound by workplace time constraints that may not be found in the work to home trips.

2. **Temporal uncertainty and Temporal expectancy**: *Temporal uncertainty* refers to how well the subject can estimate the duration of the task given previous experiences. *Temporal expectancy* refers to the accumulated previous experiences that allow the subject to generate an estimate of the duration of time for a task. The time perception research (Zakay, 1992) of psychology indicates that subjects will be more (less) accurate in estimating the time duration when temporal uncertainty is low (high). In the case of temporal expectation (Jones and Boltz, 1989, Boltz, 1993), it indicates that subjects will overestimate (underestimate) the time duration when temporal expectation is high (low).

   In transportation, psychologists define uncertainty in a dissimilar way to transportation researchers. The reason is that in transportation networks there's high interaction between travelers and also with elements (e.g. traffic signals, ramp meters) of the transportation system. Such interaction is likely to not be present in the tasks subjects must perform in psychological studies. In commute trips, there is likely to be a link between the temporal expectancy and the temporal uncertainty. The uncertainty of the trip will lead to addition or subtractions of the time elapsed in the trip due to those factors outside of the travelers control, and depending on the magnitude it is likely that those factors will enter into the memory of travelers. Thus, travelers that expect a trip of 20 minutes, and experience a trip of 40 minutes may tend to expect higher trips on the same chosen routes.

Commuters with knowledge of traffic from external sources (e.g. radio, TV) will tend to overestimate their travel time, because they are likely to have a formed a *temporal expectancy* (i.e. the received travel conditions), and have accepted a level of *temporal uncertainty* before initiating their trip. In addition, commuters' trips on bridges in Interstates will tend to underestimate their travel time, because they are likely to have high *temporal expectancy*, and also believe these bridges have low *temporal uncertainty*. Lastly, the expected travel time as indicated by the subjects is also linked to the *temporal uncertainty* and the *temporal expectancy* of a trip.

3. **Task complexity and Absorption and Attentional deployment**: *Task complexity* refers to the effort and the characteristics of the task. *Absorption and attentional deployment* refer to the focus of subjects and their understanding of the task that must be performed. In time perception research, both *task complexity* (Thomas and Weaver, 1975).and *absorption and attentional deployment* (Tellegen and Atkinson, 1974, Glicksohn and Pavell, 1992) lead to overestimation of time duration for tasks requiring more effort and more focus.

   In transportation, travelers must experience different changes in the network during their trip. There may be: hierarchical discontinuities (e.g. moving from an arterial to a freeway); the network may exhibit circuity (i.e. travelers do not just follow a straight path towards their destination); the distance traveled of the trip in a specific hierarchy (e.g. travelers may travel mostly on freeways); and others. Thus, travelers during their trip are likely to be stimulated by different changes that may require effort on their part to guarantee that they will arrive at the desired destination. Moreover, network measures are likely to be linked to *task complexity* and *absorption and attentional deployment*. Traveling through a network may require different degrees of effort on part of the commuters to keep focused. Thus, commuters traveling on paths requiring more effort should lead overestimate their travel times.

4. **Affective elements**: *Affective elements* represent emotional levels of the individuals while performing a task. The time perception research (Langer et al., 1961, Thayer and Schiff, 1975, Angrilli et al., 1997) of psychology indicates that *affective elements* lead to overestimation of time duration, because high levels of emotional stimuli makes time appear longer.

   In transportation, the emotional value of the trip is also important. Travelers that may find themselves in stressful or unpleasant situations while driving may have their perception of time affected. Congestion levels, stress levels, and fear of bridge collapse are related to the quality of the commute, and thus should be connected to *affective elements*. Commuters that are tired or stressed from the trip may have an impact on their perception. Commuters that indicate high levels of congestion in their trips will tend to overestimate their travel time, because it is believed that congested trips are unpleasant to commuters. In addition, commuters that indicate high levels of stress in their trips will tend to overestimate their travel time, because it is believed that stressful trips are unpleasant to commuters. Moreover, commuters that are experiencing fear of bridge collapse are likely to underestimate the travel times, because they are more focused on their own fears rather than quality of the travel experience.

5. **Nonlinearities**: Trips that are significantly more complex in comparison to others may lead to even higher perception error. Thus, *nonlinearities* are hypothesized to be present.

4.2.5 Statistical hypothesis testing, and goodness of fit

There are two hypothesis tests that are considered for the regression models in this study. For the nested models, the *Wald tests* are used as they only depend on the covariance matrix of the unrestricted models, and do not require estimation of the restricted models. These tests are asymptotically equivalent to the *likelihood ratio tests*. For the nonnested models, the *Akaike information criterion* (AIC), and *Bayesian information criterion* (BIC) are used in order to compare the statistical fit of similarly defined regression models (Cramer, 1986, Johnston and DiNardo, 1997, Pawitan, 2001, Greene, 2012).

The goodness of fit of the models are tested using prediction analysis (Cameron and Trivedi, 2005, 2010). The standard errors of prediction (StdP) (Cameron and Trivedi, 2005, 2010) are computed to evaluate the models accuracy. Lower values of the standard error of prediction indicate that a model fits better the data set.

4.2.6 Estimation

The linear regression models, and the logistic regression model are estimated using Maximum Likelihood methods (Cramer, 1986, Pawitan, 2001). Both set of models follow the *random effects model* for panel data (Johnston and DiNardo, 1997, Cameron and Trivedi, 2005, Wooldridge, 2010, Greene, 2012). The estimation of the models is done by maximizing their respective loglikelihood functions. For the linear regression models, the Likelihood function is of closed form. For the logistic regression model, numerical integration (Adaptive Gauss-Hermite quadrature; 30 integration points) is used as the Likelihood function is not of closed form.

The models are estimated using STATA (Cameron and Trivedi, 2010).

The Likelihood function of the linear regression models is

$$L(\beta, \sigma_\epsilon^2, \sigma_\alpha^2) = \prod_{\forall n \in \mathcal{N}} \int_{-\infty}^{\infty} \left[\prod_{\forall j \in \mathcal{J}_n} N(\tau_{jn} | f(\mathbf{x}_{jn}, \mathbf{z}_{jn}; \beta), \sigma_\epsilon^2, \alpha_n) \right] N(\alpha_n | 0, \sigma_\alpha^2) d\alpha_n \tag{4.11}$$

where

- $f(\cdot)$: Functional forms for the covariates.
- τ_{jn}: The dependent variable defined in equation 4.1.
- $\mathbf{x}_{jn}$ The vector of continuous covariates described in section 4.2.3.
- $\mathbf{z}_{jn}$ The vector of categorical covariates described in section 4.2.3.
- β: The vector of parameters to be estimated.
- $\alpha_n \sim$ i.i.d. $N(0, \sigma_\alpha^2)$ for all n
- $\epsilon_{jn} \sim$ i.i.d. $N(0, \sigma_\epsilon^2)$ for all j and n

The Likelihood function of the logistic regression model is

$$L(\beta, \sigma^2) = \prod_{\forall n \in \mathcal{N}} \int_{-\infty}^{\infty} \left[\prod_{\forall j \in \mathcal{J}_n} \left(\frac{e^{f(\mathbf{x}_{jn}, \mathbf{z}_{jn}, \alpha_n; \beta)}}{1 + e^{f(\mathbf{x}_{jn}, \mathbf{z}_{jn}, \alpha_n; \beta)}} \right)^{\delta_{jn}} \left(\frac{1}{1 + e^{f(\mathbf{x}_{jn}, \mathbf{z}_{jn}, \alpha_n; \beta)}} \right)^{1-\delta_{jn}} \right] N(\alpha_n | 0, \sigma^2) d\alpha_n \tag{4.12}$$

where

- $f(\cdot)$: Functional forms for the covariates.
- δ_{jn}: The dependent variable defined in equation 4.3.
- $\mathbf{x}_{jn}$ The vector of continuous covariates described in section 4.2.3.
- $\mathbf{z}_{jn}$ The vector of categorical covariates described in section 4.2.3.
- β: The vector of parameters to be estimated.
- $\alpha_n \sim$ i.i.d. $N(0, \sigma^2)$ for all n

4.3 Discussion and results

Tables 4.2, 4.3, and 4.4 present the estimates of the regression models in the following order: linear regression models with different functional forms (linear, quadratic, and Cobb-Douglas) for the trips with travel times underestimated (these models use only observations with $\tau_{jn} < 1$) by the subjects, and for the trips with travel times overestimated (these models use only observations with $\tau_{jn} > 1$) by the subjects; and logistic regression models (these models use observations with $\tau_{jn} < 1$, and $\tau_{jn} > 1$) to predict whether the travel time of a trip will be overestimated or underestimated by the subjects. Furthermore, it should be noted that the direction of the gradient of regressors (e.g. $\nabla\tau_{jn}(\mathbf{x}_{jn};\hat{\beta})$) in the functions with the estimated coefficients ($\tau_{jn} = f(\mathbf{x}_{jn};\hat{\beta})$) has different interpretation for the underestimated trips ($\tau_{jn} < 1$), and for the overestimated trips ($\tau_{jn} > 1$). In other words, positive values of first order derivative of a regressor in $\tau_{jn}(\mathbf{x}_{jn};\hat{\beta})$ increase the likelihood of $\tau_{jn} \to 1$ (*ceteris paribus*) for linear regression models for *underestimated trips.* In contrast, positive values of first order derivative of a regressor in $\tau_{jn}(\mathbf{x}_{jn};\hat{\beta})$ decrease the likelihood of $\tau_{jn} \to 1$ (*ceteris paribus*) for linear regression models for *overestimated trips.* In the case of logistic regression models, positive values of first order derivative of a regressor in $f(\mathbf{x}_{jn};\hat{\beta})$ increase the likelihood of $Prob[\delta_{jn} = 1] \to 1$ (overestimation of travel times in commute trips; *ceteris paribus*). Readers should refer to section 4.2 for details.

Each of the hypotheses of this study (see section 4.2.4) are discussed subsequently within the results of each of the groups of regression models mentioned in the previous paragraph. The results are discussed within a *ceteris paribus* context.

The variables representing *arrival flexibility* are only statistically significant at least 10% in the linear regression models for trips with travel times underestimated. Commute trips (with travel times underestimated) of subjects with higher *arrival flexibility* to work are more likely to further underestimate their travel time ($\tau_{jn} \to 0$) in comparison to other commute trips (with travel times underestimated) of subjects. The variable representing *type of trip* (home to work, and work to home) is statistically significant at least 5% in the logistic regression models, and the linear regression models for trips with travel times underestimated. Commute trips from home to work (with travel times underestimated) of subjects are less likely to further underestimate their travel time ($\tau_{jn} \rightarrow 1$) in comparison to commute trips from work to home (with travel times underestimated) of subjects. In addition, commute trips from home to work of subjects are more likely to overestimate their travel time ($Prob[\delta_{jn} = 1] \rightarrow 1$) in comparison to commute trips from work to home of subjects. However, these variables are not statistically significant in the linear regression models for trips with travel times overestimated. These results may be explained by the link these variables are likely to have with **temporal relevance** as suggested previously.

The variable representing *traffic information* is only statistically significant at least 10% in the linear regression models for trips with travel times overestimated, and the logistic regression models. Commute trips (with travel times overestimated) of subjects using *traffic information* are less likely to further overestimate their travel time ($\tau_{jn} \rightarrow 1$) in comparison to commute trips (with travel times overestimated) of subjects using no traffic information. In addition, commute trips of subjects using *traffic information* are more likely to overestimate their travel time ($Prob[\delta_{jn} = 1] \rightarrow 1$) in comparison to commute trips of subjects using no traffic information.

The variable representing *trips on Interstate bridges* is only statistically significant at least 5% in the linear regression models for trips with travel times underestimated, and the logistic regression models. Commute trips (with travel times underestimated) of subjects traveling on *Interstate bridges* are less likely to further underestimate their travel time ($\tau_{jn} \rightarrow 1$) in comparison to commute trips (with travel times underestimated) of subjects on other bridges. In addition, commute trips of subjects traveling on *Interstate bridges* are less likely to overestimate their travel time ($Prob[\delta_{jn} = 1] \rightarrow 0$) in comparison to commute trips of subjects on other bridges.

Furthermore, the presence of the *expected travel time* of their trips as indicated by the subjects is statistically significant at least 5% in all models. However, the *expectation of travel times* is necessarily linked to other unknown variables not included as regressors, and thus it is likely to be correlated with the ϵ_{jn} terms (i.e. endogenous) in the regression models. The sign of the *expected travel time* variable is negative in all the models. In general terms, this indicates that higher expectation of travel times is: more likely to further underestimate travel times in trips with underestimated travel times ($\tau_{jn} \rightarrow 0$); less likely to further overestimate travel times in trips with overestimated travel times ($\tau_{jn} \rightarrow 1$); and less likely to lead to overestimation of travel times in trips ($Prob[\delta_{jn} = 1] \rightarrow 0$). The direction of the *expected travel time* variable, as mentioned previously, must be interpreted carefully. It must be remembered that subjects' *expectation of travel time* may be influenced by unknown variables including but not limited to: past experiences; and presence of anchors (e.g. signals that provide confidence in an uncertain environment).

The results of *traffic information*, *trips on Interstate bridges*, and *expected travel time* may be explained by the link these variables are likely to have with **temporal uncertainty and temporal expectancy** as suggested previously.

The variables representing *relative discontinuity* are only statistically significant at least 5% in the linear regression models for trips with travel times overestimated. Commute trips (with travel times overestimated) of subjects with higher *relative discontinuity* are more likely to further overestimate their travel time ($\tau_{jn} \rightarrow \infty$) in comparison to other commute trips (with travel times overestimated) of subjects. In addition, the *relative discontinuity* exhibits statistical significant of at least 5% nonlinearities in the quadratic functional form, and the Cobb-Douglas functional form for commute trips (with travel times overestimated) of subjects. In the quadratic functional form, increases in further overestimation of travel times by increases in *relative discontinuity* ($\frac{df(\cdot)}{dx_{jn}} = \gamma_1 - 2\gamma_2 x_{jn}$; $\gamma_1 > 2\gamma_2 x_{jn}$) are smaller the more it increases up until a point ($\frac{df(\cdot)}{dx_{jn}} = 0; \gamma_1 = 2\gamma_2 x_{jn}$), and after such point increases in *relative discontinuity* leads to decreases in further overestimation of travel times ($\frac{df(\cdot)}{dx_{jn}} = \gamma_1 - 2\gamma_2 x_{jn}$; $\gamma_1 < 2\gamma_2 x_{jn}$). In contrast, the increases of further overestimation of travel times by increases in *relative discontinuity* leads to decreases in further overestimation in the Cobb-Douglas functional form ($\frac{df(\cdot)}{dx_{jn}} = \gamma(x_{jn})^{\gamma-1}$; $\gamma < 0$). It should be noted that γ are parameters representing the corresponding $\hat{\beta}$ estimates from the table 5.1. Thus, both functional forms indicate decreases in the further overestimation of travel times, but the quadratic functional form accommodates increases of further overestimation up until a point.

The variables representing *proportion of limited access roads* are only statistically significant at least 5% in the linear regression models for trips with travel times overestimated, and the logistic regression models. Commute trips (with travel times overestimated) of subjects with higher *proportion of limited access roads* are less likely to further overestimate their travel time ($\tau_{jn} \rightarrow 1$) in comparison to other commute trips (with travel times overestimated) of subjects. In addition, commute trips of subjects with higher *proportion of limited access roads* are less likely to overestimate their travel time ($Prob[\delta_{jn} = 1] \rightarrow 0$) in comparison to other commute trips of subjects. Nonlinearities are not found statistically significant even at 10% for the *proportion of limited access roads*.

The variables representing *proportion of signalized arterials* are only statistically significant at least 10% in the linear regression models for trips with travel times overestimated, and for trips with travel times underestimated, and in the logistic regression models. Commute trips (with travel times overestimated) of subjects with higher *proportion of signalized arterials* are more likely to further overestimate their travel time ($\tau_{jn} \rightarrow \infty$) in comparison to other commute trips (with travel times overestimated) of subjects. Also, commute trips (with travel times underestimated) of subjects with higher *proportion of signalized arterials* are less likely to further underestimate their travel time ($\tau_{jn} \rightarrow 1$) in comparison to other commute trips (with travel times underestimated) of subjects. In addition, commute trips of subjects with higher *proportion of signalized arterials* are more likely to overestimate their travel time ($Prob[\delta_{jn} = 1] \rightarrow 1$) in comparison to other commute trips of subjects. Moreover, the *proportion of signalized arterials* exhibits statistical significant of at least 5% nonlinearities in the quadratic functional form, and the Cobb-Douglas functional form for commute trips with travel times overestimated, and for commute trips with travel times underestimated. Also, there's nonlinearity statistically significant at 5% through a quadratic term in the logistic regression models. In the models, both the functional forms (quadratic and Cobb-Douglas) exhibit concavity properties: $\frac{df(\cdot)}{dx_{jn}} = \gamma_1 - 2\gamma_2 x_{jn}$, $\gamma_1 > 0$, $\gamma_2 > 0$ (quadratic); $\frac{df(\cdot)}{dx_{jn}} = \gamma(x_{jn})^{\gamma-1}$, $\gamma > 0$ (Cobb-Douglas). Therefore, increases in *proportion of signalized arterials* lead to smaller increases in τ_n^i (linear regression models)/$Prob[\delta_{jn} = 1]$ (logistic regression models) as the *proportion of signalized arterials* further increases. In addition, decreases in τ_{jn} (linear regression models)/$Prob[\delta_{jn} = 1]$ (logistic regression models) by increases in *proportion of signalized arterials* ($\frac{df(\cdot)}{dx_{jn}} = \gamma_1 - 2\gamma_2 x_{jn}$; $\gamma_1 < 2\gamma_2 x_{jn}$) will proceed after a point ($\frac{df(\cdot)}{dx_{jn}} = 0; \gamma_1 = 2\gamma_2 x_{jn})$) in the quadratic functional forms. It should be noted that γ are parameters representing the corresponding $\hat{\beta}$ estimates from the table 5.1.

The variables representing *circuity* are only statistically significant at least 5% in the linear regression models for trips with travel times overestimated, and the logistic regression models. Commute trips (with travel times overestimated) of subjects with higher *circuity* are more likely to further overestimate their travel time ($\tau_{jn} \rightarrow \infty$) in comparison to other commute trips (with travel times overestimated) of subjects. In addition, commute trips of subjects with higher *circuity* are more likely to overestimate their travel time ($Prob[\delta_{jn} = 1] \rightarrow 1$) in comparison to other commute trips of subjects. Moreover, the *circuity* exhibits statistical significant of at least 5% nonlinearities in the quadratic functional form for the logistic regression models, and the Cobb-Douglas functional form for commute trips with travel times overestimated. In the Cobb-Douglas functional form, increases of further overestimation of travel times by increases in *circuity* leads to smaller increases in further overestimation as the *circuity* continues to increase ($\frac{df(\cdot)}{dx_{jn}} = \gamma(x_{jn})^{\gamma-1}$; $\gamma > 0$). In the quadratic functional form, increases in $Prob[\delta_n^i = 1]$ by increases in *circuity* ($\frac{df(\cdot)}{dx_{jn}} = \gamma_1 - 2\gamma_2 x_{jn}$; $\gamma_1 > 2\gamma_2 x_{jn}$) are smaller the more it increases up until a point ($\frac{df(\cdot)}{dx_{jn}} = 0; \gamma_1 = 2\gamma_2 x_{jn}$), and after such point increases in *circuity* leads to decreases in $Prob[\delta_{jn} = 1]$ ($\frac{df(\cdot)}{dx_{jn}} = \gamma_1 - 2\gamma_2 x_n^i$; $\gamma_1 < 2\gamma_2 x_{jn}$). It should be noted that γ are parameters representing the corresponding $\hat{\beta}$ estimates from the table 5.1.

The results of *relative discontinuity*, *proportion of limited access roads*, *proportion of signalized arterials*, and *circuity* may be explained by the link these variables are likely to have with **task complexity and absorption and attentional deployment** as suggested previously.

The variables representing *congestion levels* are only statistically significant at least 5% in the linear regression models for trips with travel times underestimated, and in the logistic regression models. Commute trips (with travel times underestimated) of subjects with higher *congestion levels* are more likely to further underestimate their travel time ($\tau_{jn} \rightarrow 0$) in comparison to other commute trips (with travel times underestimated) of subjects. In addition, commute trips of subjects with higher *congestion levels* are less likely to overestimate their travel time ($Prob[\delta_{jn} = 1] \rightarrow 0$) in comparison to other commute trips of subjects. The variables representing *stress levels* are only statistically significant at least 10% in the linear regression models for trips with travel times underestimated. Commute trips (with travel times underestimated) of subjects with higher *stress levels* are more likely to further underestimate their travel time ($\tau_{jn} \rightarrow 0$) in comparison to other commute trips (with travel times underestimated) of subjects. The variable representing *fear of driving on bridges due to a previous bridge collapse* is not statistically significant even at 10% in any of the regression models.

The results of *congestion levels* and *stress levels* may not be explained by the link these variables are likely to have with **affective elements** as suggested previously. In fact, It is unknown whether the subjects' understanding of the abstract situation matches the authors intention (i.e. high *congestion levels*, and high *stress levels* are unpleasant) of the abstract situation. Subjects may tolerate high *congestion levels* in their trips, because of their continuous recurrence. Similarly, subjects may tolerate high *stress levels* in their trips, because of their continuous recurrence. In addition, subjects' attitudes (e.g. optimism) toward *congestion levels*, and *stress levels* may dominate.

Socio-demographic variables (*gender*, and *income*) were added to control for the *observed heterogeneity* of the subjects. *Gender* was not found statistically significant even at 10% for any of the models. In contrast, the variables related to *income* are found statistically significant at 5% in the linear regression models for trips with overestimated travel times. The sign is negative, and thus indicating that commute trips by subjects with higher *income* are less likely to further overestimate their travel time ($\tau_{jn} \rightarrow 1$) in comparison to other commute trips (with travel times overestimated) of subjects. These variables require further research to identify the reason behind them. It is hypothesized that perhaps the type of jobs may play a role. Furthermore, the variables representing *random effect* are statistically significant at 1% in all the models. This indicates that there are unobservable variables (i.e. unobserved heterogeneity) attributed to each subjects influencing their own perception of travel time in their trips. This is expected as perception is likely to vary across subjects.

Lastly, the goodness of fit of the models indicate that the linear functional form for the linear regression models (for trips with travel times underestimated) fits better to the data set than the quadratic functional form, and the Cobb-Douglas functional form. In contrast, the Cobb-Douglas functional form for the linear regression models (for trips with travel times overestimated) fits better to the data set than the quadratic functional form, and the linear functional form. This is not surprising as the linear regression models for trips with travel times overestimated exhibit more statistically significant (at least 10%) nonlinearities than the linear regression models for trips with travel times underestimated. Moreover, it should also be noted that the final loglikelihood of some models is positive, but this should not cause concern as loglikelihoods can be positive in linear regression models with Normal disturbances. In addition, table 4.1 presents the correlation matrix of the regressors. The correlations are low to modest across regressors.

4.4 Conclusion

In a similar vein, this study follows previous research (Parthasarathi, 2011, Parthasarathi et al., 2012), and extends it by considering additional factors (e.g. arrival flexibility, access to traffic information; see section 4.2.3) beyond those related to road network structure, and also by ascertaining nonlinearities in those factors related to road network structure. Furthermore, these factors are included in econometric models to study their influence on travel time perception, and also ascertains which these factors lead to overestimation or underestimation of travel times. These econometric models, described in section 4.2, are fitted using three functional forms (linear, quadratic, and Cobb-Douglas) on data collected of commuters recruited from a previous research study in the Minneapolis-St. Paul region (Zhu, 2010, Carrion and Levinson, 2012a). This data (surveys, and Global Positioning System [GPS] points) consists of work trips (from home to work, and from work to home) of subjects. For these work trips, the subjects' self-reported travel times, and the subjects' travel times measured by GPS devices were collected. In addition, the factors are related into four categories based on time perception research in psychology: temporal relevance; temporal uncertainty, and temporal expectancies; task complexity, and absorption, and attentional deployment; and affective elements. This allows the factors to be compared to similar research of time perception in psychology.

These results continue to highlight the need to further study the perception of travel time, and to acknowledge its influence in travelers' decisions. Thus, the modeling of travel decisions must account for perception error. This is an important research topic as many analyses (e.g. economic, planning), and models (e.g. economic, traffic) in transportation continue to ignore that travelers are executing decisions according to their own divergent views of the actual travel time distributions.

Table 4.1: Correlation matrix of regressors

Variables	**Arrival Flexibility:** Within 20 minutes of start time	**Arrival Flexibility:** Within 60 minutes of start time	**Arrival Flexibility:** At any time beyond start time	**Type of trip:** home to work	**Expected travel time**	**Traffic information**	**Interstate Bridge**	**Relative discontinuity**	**Proportion of limited access roads**	**Proportion of signalized arterials**	**Circuity**	**Congestion level:** Medium	**Congestion level:** High	Stress level: Medium	**Stress level:** High	**Fear of driving near I-35W bridge**	**Gender:** Male	**Income:** [$50,000, $99,999]	**Income:** [$100,000, ∞+)
Arrival Flexibility: Within 20 minutes of start time	1.0000																		
Arrival Flexibility: Within 60 minutes of start time	-0.3451	1.0000																	
Arrival Flexibility: At any time beyond start time	-0.4595	-0.2484	1.0000																
Type of trip: home to work	0.1822	-0.0699	-0.3691	1.000															
Expected travel time	0.0095	0.0020	0.0621	-0.0813	1.0000														
Traffic information	0.0042	-0.1274	-0.0751	0.1606	-0.0140	1.0000													
Interstate Bridge	0.0923	0.0191	-0.0866	0.0196	-0.0268	-0.0137	1.0000												
Relative discontinuity	0.0034	-0.0130	-0.0651	0.0622	0.1594	0.1056	0.0474	1.0000											
Proportion of limited access roads	0.0786	-0.0155	-0.1137	0.0933	0.2189	0.1018	0.1946	0.7239	1.0000										
Proportion of signalized arterials	-0.0397	-0.1035	0.0677	0.0678	0.2271	0.0523	-0.1419	0.3584	0.2141	1.0000									
Circuity	-0.0579	0.0637	-0.0905	0.0617	0.0699	0.1251	0.0825	0.1235	0.1381	0.0597	1.0000								
Congestion level: Medium	-0.0159	-0.0199	-0.0221	-0.1400	0.1689	-0.0446	-0.0274	-0.0209	-0.1054	0.1408	0.0160	1.0000							
Congestion level: High	-0.0555	0.0083	0.0228	-0.0373	-0.0026	-0.0056	-0.0250	-0.0517	-0.0608	-0.0692	-0.0241	-0.2202	1.0000						
Stress level: Medium	0.0631	-0.0236	-0.0815	-0.0448	0.0996	-0.0172	-0.0403	-0.0815	-0.1773	0.1285	-0.0574	0.6219	0.0910	1.0000					
Stress level: High	-0.0389	0.0003	-0.0132	0.0114	-0.0149	-0.0514	-0.0465	-0.0793	-0.0769	-0.0684	0.0039	-0.1006	0.5234	-0.1247	1.0000				
Fear of driving near I-35W bridge	-0.0134	0.0163	-0.0183	-0.0253	0.0475	0.1634	-0.0155	0.0525	-0.0253	0.0470	0.2574	0.0409	-0.0218	0.0786	0.0013	1.0000			
Gender: Male	0.0810	0.0592	0.0299	0.0009	-0.1191	-0.0074	0.0149	-0.0313	-0.0769	-0.0312	-0.0490	-0.0307	-0.0878	-0.0735	-0.0691	-0.0745	1.0000		
Income: [$50,000, $99,999]	-0.0590	0.0609	0.1092	-0.0644	0.0404	0.0262	-0.1901	0.0625	-0.0955	0.2802	-0.0251	0.1312	0.0044	0.1853	-0.0547	-0.0294	0.0761	1.0000	
Income: [$100,000, ∞+)	0.0034	0.0394	-0.0260	0.0373	0.0835	-0.1093	0.0983	-0.0292	0.1009	-0.1453	-0.0883	-0.0356	0.0108	-0.0958	0.0420	-0.0163	0.0269	-0.7022	1.0000

Table 4.2: Linear Regression Models - Underestimated

Variables	Linear - Underestimated		Quadratic - Underestimated		Cobb-Douglas - Underestimated		
Regression model: **Dependent variable:**	Linear regression model τ_{jn}; Equation 4.1		Linear regression model τ_{jn}; Equation 4.1		Linear regression model τ_{jn}; Equation 4.1		**Hypotheses**[a]**:**
	Estimates (T-Stats)		**Estimates (T-Stats)**		**Estimates (T-Stats)**		
Arrival Flexibility: Within 20 minutes of start time (1 = In; 0 = Out)	-0.032 (-1.52)	-0.040 (-1.72) *	-0.029 (-1.44)	-0.040 (-1.71) *	-0.025 (-0.74)	-0.052 (-1.30)	**-S**
Arrival Flexibility: Within 60 minutes of start time (1 = In; 0 = Out)	-0.062 (-2.45) **	-0.075 (-2.68) **	-0.059 (02.43) **	-0.077 (-2.74) **	-0.064 (-1.58)	-0.096 (-1.98) **	**-S**
Arrival Flexibility: At any time beyond start time (1 = In; 0 = Out)	-0.0719 (-3.07) **	-0.099 (-3.83) ***	-0.068 (-2.97) **	-0.100 (-3.86) ***	-0.095 (-2.53) **	-0.16 (-3.57) ***	**-S**
Type of trip (1 = from home to work; 0 = from work to home)	0.021 (1.55)	0.039 (2.58) **	0.012 (0.88)	0.038 (2.48) **	-0.00095 (-0.04)	0.065 (2.49) **	**+S**
Expected travel time	-0.0057 (-11.70) ***		-0.010 (-7.69) ***				
Expected travel time - squared/logarithmed			0.000033 (3.68) ***		-0.59 (-15.37) ***		
Traffic Information (1 = received information; 0 = did not receive information)	-0.025 (-1.02)	-0.038 (-1.37)	-0.017 (-0.68)	-0.035 (-1.26)	-0.024 (-0.58)	-0.062 (-1.29)	+S
Interstate Bridge (1 = used an Interstate bridge; 0 = did not use an Interstate bridge)	-0.036 (-2.56) **	-0.028 (-1.81) *	-0.035 (-2.58) **	-0.027 (-1.74) *	-0.070 (-3.12) **	-0.047 (-1.75)	**-S**
Relative discontinuity	0.0024 (0.13)	0.0051 (0.25)	-0.079 (-1.46)	-0.043 (-0.76)			+S
Relative discontinuity - squared/logarithmed			0.01 (1.49)	0.01 (0.85)	-0.054 (-2.25) **	-0.0146 (-0.61)	
Proportion of limited access roads	0.016 (0.34)	-0.009 (-0.19)	0.28 (1.30)	0.033 (0.14)			-S
Proportion of limited access roads - squared/logarithmed			-0.27 (-1.16)	-0.041 (-0.16)	0.0058 (0.48)	-0.0052 (-0.66)	
Proportion of signalized arterials	0.26 (4.37) ***	0.166 (2.83) **	0.54 (1.75) *	0.46 (1.35)			**+S**
Proportion of signalized arterials - squared/logarithmed			-0.44 (-0.95)	-0.49 (-0.94)	0.073 (3.44) **	0.0296 (1.28)	
Circuity	0.013 (0.63)	-0.017 (-0.81)	0.079 (0.82)	0.012 (0.13)			+S
Circuity - squared/logarithmed			-0.011 (-0.74)	-0.006 (-0.38)	0.076 (0.84)	-0.097 (-1.10)	
Congestion level Medium congestion level (1 = In; 0 = Out)	-0.054 (-3.04) ***	-0.063 (-3.13) **	-0.056 (-3.19) ***	-0.065 (-3.22) **	-0.055 (-1.94)	-0.083 (-2.40) **	*+S*
Congestion level High congestion level (1 = In; 0 = Out)	-0.189 (-5.64) ***	-0.173 (-4.52) ***	-0.191 (-5.87) ***	-0.175 (-4.59) ***	-0.27 (-5.13)	-0.16 (-3.75) ***	*+S*
Stress level Medium stress level (1 = In; 0 = Out)	-0.029 (-1.63)	-0.023 (-1.15)	-0.032 (-1.83) *	-0.025 (-1.23)	-0.040 (-1.41)	-0.037 (-1.06)	*+S*
Stress level High stress level (1 = In; 0 = Out)	-0.088 (-1.82) *	-0.089 (-1.60)	-0.093 (-1.98) **	-0.089 (-1.61)	-0.17 (-2.29) **	-0.16 (-1.65) *	*+S*
Fear of driving on the I-35W bridge and other bridges in the vicinity (1 = In; 0 = Out)	0.044 (1.47)	0.037 (1.38)	0.050 (1.42)	0.035 (1.24)	0.097 (1.35)	0.061 (1.36)	-S
Gender [1= Male; 0 = Female]	0.014 (0.49)	0.040 (1.53)	0.013 (0.37)	0.042 (1.55)	-0.022 (-0.32)	0.056 (1.30)	
Income [$50,000, $99,999] (1 = In; 0 = Out)	0.029 (0.68)	-0.0024 (-0.06)	0.028 (0.58)	-0.0078 (-0.20)	0.073 (0.75)	-0.027 (-0.42)	
Income [$100,000, ∞+) (1 = In; 0 = Out)	0.037 (0.79)	0.0079 (0.19)	0.033 (0.62)	-0.0024 (-0.06)	0.11 (1.02)	0.0031 (0.04)	
Intercept/Scale constant	0.94 (16.22) ***	0.84 (15.11) ***	0.98 (8.25) ***	0.81 (7.19) ***	1.81 (11.05) ***	-0.0798 (-0.86)	
Dispersion parameter σ_ϵ	0.124 (30.56) ***	0.143 (30.84) ***	0.12 (30.12) ***	0.14 (30.77) ***	0.19 (30.15) ***	0.25 (30.79) ***	
Random effect σ_α	0.093 (8.07) ***	0.075 (7.26) ***	0.11 (7.94) ***	0.078 (7.26) ***	0.24 (8.66) ***	0.12 (6.71) ***	
Intercept Log-Likelihood $ll_{A\hat{S}C}$	194.151	194.151	194.151	194.151	-87.790	-87.790	
Final Log-Likelihood $ll_{\hat{\beta}}$	306.423	245.667	316.827	247.096	49.254	-43.537	
Akaike Information Criterion AIC	-568.846	-449.335	-579.654	-444.192	-54.50895	39.41117	
Bayesian Information Criterion BIC	-474.926	-359.684	-464.389	-337.465	129.074	218.725	
Standard Error of Prediction $StdP$	0.0416163	0.0396277	0.0492628	0.0435538	0.0939007	0.0679951	
Number of observations	528	528	528	528	528	528	
Number of subjects	60	60	60	60	60	60	

* is 10% significance level, ** is 5% significance level, *** is 1% significance level

[a] **Bold** corroborates hypothesis, *Italic* refutes hypothesis, Normal does not corroborates or refutes hypothesis

See the section 4.2 for details on the econometric models.

Table 4.3: Linear Regression Models - Overestimated

Variables	Linear - Overestimated		Quadratic - Overestimated		Cobb-Douglas - Overestimated		
Regression model: Dependent variable:	Linear regression model τ_{jn}; Equation 4.1		Linear regression model τ_{jn}; Equation 4.1		Linear regression model τ_{jn}; Equation 4.1		**Hypotheses**[a]**:**
	Estimates (T-Stats)		**Estimates (T-Stats)**		**Estimates (T-Stats)**		
Arrival Flexibility: **Within 20 minutes of start time** **(1 = In; 0 = Out)**	0.078 (0.46)	0.19 (1.09)	0.076 (0.48)	0.22 (1.32)	0.059 (1.07)	0.086 (1.44)	-S
Arrival Flexibility: **Within 60 minutes of start time** **(1 = In; 0 = Out)**	0.27 (1.47)	0.29 (1.47)	0.22 (1.27)	0.26 (1.36)	0.061 (1.00)	0.071 (1.03)	-S
Arrival Flexibility: **At any time beyond start time** **(1 = In; 0 = Out)**	0.014 (0.07)	0.19 (0.95)	-0.039 (-0.22)	0.17 (0.88)	-0.0056 (-0.09)	0.067 (0.96)	-S
Type of trip **(1 = from home to work;** **0 = from work to home)**	-0.099 (-0.82)	-0.055 (-0.42)	-0.105 (-0.93)	-0.072 (-0.56)	-0.049 (-1.24)	-0.04 (-0.84)	+S
Expected travel time	-0.051 (-6.09) ***		-0.107 (-3.23) **				
Expected travel time - squared/logarithmed				0.00089 (1.59)		-0.56 (-8.96) ***	
Traffic Information **(1 = received information;** **0 = did not receive information)**	-0.103 (-0.61)	-0.15 (-0.83)	-0.062 (-0.40)	-0.19 (-1.14)	-0.10 (-1.88) *	-0.16 (-2.63) **	**+S**
Interstate Bridge **(1 = used an Interstate bridge;** **0 = did not use an Interstate bridge)**	0.04 (0.33)	-0.022 (-0.19)	-0.011 (-0.12)	-0.048 (-0.43)	-0.049 (-1.41)	-0.052 (-1.31)	-S
Relative discontinuity	0.12 (1.24)	0.11 (0.99)	0.91 (2.35) **	1.10 (2.57) **			**+S**
Relative discontinuity - squared/logarithmed			-0.158 (-2.15) **	-0.198 (-2.43) **	-0.09 (-2.59) **	-0.085 (-2.22) **	
Proportion of limited access roads	-1.58 (-3.42) **	-1.09 (-2.97) **	-1.20 (-0.63)	-1.69 (-0.89)			**-S**
Proportion of limited access roads - squared/logarithmed			-2.96 (-1.39)	-1.63 (-0.76)	0.0018 (0.10)	0.010 (0.71)	
Proportion of signalized arterials	1.21 (2.56) **	0.61 (1.45)	6.26 (2.29) **	4.69 (1.60)			**+S**
Proportion of signalized arterials - squared/logarithmed			-12 (-2.66) **	-9.55 (-2.03) **	0.11 (3.20) **	0.083 (2.20) **	
Circuity	1.38 (6.54) ***	0.95 (5.61) ***	1.06 (1.73) **	1.62 (3.13) **			**+S**
Circuity - squared/logarithmed			0.061 (0.50)	-0.13 (-1.45)	0.67 (5.72) ***	0.56 (4.70) ***	
Congestion level **Medium congestion level** **(1 = In; 0 = Out)**	-0.15 (-0.93)	-0.24 (-1.35)	-0.066 (-0.42)	-0.25 (-1.45)	0.0041 (0.08)	-0.049 (-0.80)	+S
Congestion level **High congestion level** **(1 = In; 0 = Out)**	0.0056 (0.01)	-0.14 (-0.33)	0.029 (0.08)	-0.20 (-0.48)	0.031 (0.24)	0.0019 (0.01)	+S
Stress level **Medium stress level** **(1 = In; 0 = Out)**	0.11 (0.65)	0.079 (0.42)	0.0019 (0.01)	0.037 (0.20)	-0.0112 (-0.20)	0.00088 (0.01)	+S
Stress level **High stress level** **(1 = In; 0 = Out)**	-0.40 (-0.94)	-0.27 (-0.57)	-0.47 (-1.16)	-0.25 (-0.54)	-0.21 (-1.54)	-0.19 (-1.23)	+S
Fear of driving on the I-35W bridge **and other bridges in the vicinity** **(1 = In; 0 = Out)**	-0.34 (-1.27)	-0.097 (-0.50)	-0.49 (-1.62)	-0.21 (-1.08)	-0.015 (-0.18)	0.012 (0.16)	-S
Gender **[1= Male; 0 = Female]**	-0.08 (-0.33)	0.17 (0.92)	0.0159 (0.06)	0.166 (0.91)	(-0.05) (-0.62)	0.05 (0.71)	
Income **[$\$50,000, \$99,999$]** **(1 = In; 0 = Out)**	-0.53 (-1.38)	-0.60 (-2.21) **	-0.77 (-1.72) **	-0.75 (-2.67) **	(-0.076) (-0.59)	-0.13 (-1.21)	
Income **[$\$100,000, \infty +$)** **(1 = In; 0 = Out)**	-0.21 (-0.50)	-0.60 (-2.16) **	-0.32 (-0.67)	-0.73 (-2.53)	-0.0086 (-0.06)	-0.183 (-1.64)	
Intercept/Scale constant	1.43 (3.01) **	0.72 (1.99) **	2.72 (2.94) **	0.20 (0.32)	2.13 (8.66) ***	0.37 (2.72) **	
Dispersion parameter σ_ϵ	0.76 (19.74) ***	0.88 (21.55) ***	0.70 (19.25) ***	0.84 (21.90) ***	0.25 (22.58) ***	0.30 (22.92) ***	
Random effect σ_α	0.82 (4.97) ***	0.48 (4.10) ***	0.98 (5.14) ***	0.50 (4.66) ***	0.28 (7.42) ***	0.20 (6.13) ***	
Intercept Log-Likelihood ll_{ASC}	-490.0183	-490.0183	-490.0183	-490.0183	-130.1009	-130.1009	
Final Log-Likelihood $ll_{\hat{\beta}}$	-441.6079	-460.3384	-425.785	-449.039	-67.8643	-105.7193	
Akaike Information Criterion AIC	927.2158	962.6767	905.5692	948.0778	179.7286	253.4387	
Bayesian Information Criterion BIC	1011.258	1042.898	1008.711	1043.58	263.7704	333.6604	
Standard Error of Prediction $StdP$	0.3704983	0.2941868	0.4217648	0.3103562	0.122135	0.1095125	
Number of observations	337	337	337	337	337	337	
Number of subjects	60	60	60	60	60	60	

* is 10% significance level, ** is 5% significance level, *** is 1% significance level

[a] **Bold** corroborates hypothesis, *Italic* refutes hypothesis, Normal does not corroborates or refutes hypothesis

See the section 4.2 for details on the econometric models.

Table 4.4: Logistic Regression Models

Variables	Logistic regression		
Regression model:	Logistic regression model		
Dependent variable:	δ_{jn}; Equation 4.3		**Hypotheses[a]:**
	Estimates (T-Stats)		
Arrival Flexibility: Within 20 minutes of start time (1 = In; 0 = Out)	0.096 (0.29)	-0.12 (-0.39)	-S
Arrival Flexibility: Within 60 minutes of start time (1 = In; 0 = Out)	0.33 (0.83)	0.025 (0.07)	-S
Arrival Flexibility: At any time beyond start time (1 = In; 0 = Out)	-0.040 (-0.11)	-0.17 (-0.48)	-S
Type of trip (1 = from home to work; 0 = from work to home)	0.35 (1.49)	0.59 (2.78) **	**+S**
Expected travel time	-0.134 (-6.53) ***		
Traffic Information (1 = received information; 0 = did not receive information)	0.65 (1.84) *	0.48 (1.47)	**+S**
Interstate Bridge (1 = used an Interstate bridge; 0 = did not use an Interstate bridge)	-0.59 (-2.70) **	-0.47 (-2.35) **	**-S**
Relative discontinuity	0.33 (1.18)	0.27 (1.02)	+S
Proportion of limited access roads	-1.53 (-1.62)	-1.81 (-2.28) **	**-S**
Proportion of signalized arterials	9.06 (2.56) **	5.81 (1.92) **	**+S**
Proportion of signalized arterials - squared/logarithmed	-10.54 (-1.76) *	-7.75 (-1.46)	
Circuity	4.86 (3.28) **	5.07 (4.17) ***	**+S**
Circuity - squared/logarithmed	-0.55 (-2.44) **	-0.71 (-3.61) ***	
Congestion level Medium congestion level (1 = In; 0 = Out)	-0.73 (-2.39) **	-1.06 (-3.83) ***	*+S*
Congestion level High congestion level (1 = In; 0 = Out)	-1.96 (-3.09) **	-2.24 (-3.65)	*+S*
Stress level Medium stress level (1 = In; 0 = Out)	-0.105 (-0.32)	-0.024 (-0.08)	+S
Stress level High stress level (1 = In; 0 = Out)	0.589 (0.73)	0.94 (1.26)	+S
Fear of driving on the I-35W bridge and other bridges in the vicinity (1 = In; 0 = Out)	-0.37 (-0.68)	-0.45 (-1.10)	-S
Gender [1= Male; 0 = Female]	-0.42 (-0.79)	-0.063 (-0.16)	
Income [$\mathbf{\$50,000,\$99,999}$] **(1 = In; 0 = Out)**	0.84 (1.07)	0.32 (0.54)	
Income [$\mathbf{\$100,000,\infty+}$) **(1 = In; 0 = Out)**	1.42 (1.66)	0.62 (0.98)	
Intercept	03.37 (-1.85) *	-6.07 (-4.12) ***	
Random effect σ	1.76 (6.07) ***	1.25 (6.36) ***	
Intercept Log-Likelihood $ll_{A\hat{S}C}$	-578.31024	-578.31024	
Final Log-Likelihood $ll_{\hat{\beta}}$	-440.6271	-480.2445	
Likelihood ratio index ρ^2	0.23807827	0.16957293	
Akaike Information Criterion AIC	927.2543	1004.489	
Bayesian Information Criterion BIC	1036.797	1109.269	
Number of observations	865	865	
Number of subjects	64	64	

* is 10% significance level, ** is 5% significance level, *** is 1% significance level

[a] **Bold** corroborates hypothesis, *Italic* refutes hypothesis, Normal does not corroborates or refutes hypothesis

See the section 4.2 for details on the econometric models.

Chapter 5

Uncovering the influence of commuters' perception on the reliability ratio

5.1 Introduction

Two of the most important values obtained from travel demand studies are the *value of travel time savings* (VOT), and the *value of travel time reliability* (VOR). The first refers to the marginal rate of substitution between travel cost and reductions in travel time (i.e. savings). The second refers to the marginal rate of substitution between travel cost, and increases in the predictability (i.e reducing the variability) of travel time. The ratio between these values is known as the *reliability ratio* (RR). This ratio refers to the marginal rate of substitution between reductions in travel time (i.e. savings), and increases in the predictability of travel time (i.e. reduce variability). The *value of travel time savings* has a long established history along with a firm theoretical background (i.e. the time allocation models), and many empirical estimates calculated by practitioners and researchers. Common resources on the theoretical foundations and (brief) empirical discussions are: (Bruzelius, 1979, Jara-Diaz, 2000, 2007, Small and Verhoef, 2007). Readers may also refer to detailed reviews of empirical estimates of VOT such as: (Wardman, 1998, 2001, 2004, Zamparini and Reggiani, 2007a,b, Shires and de Jong, 2009, Abrantes and Wardman, 2011). The *value of travel time* *reliability* traces its recent (quantitative) history back to the 1980s and 1990s with important contributions such as: (Jackson and Jucker, 1982, Small, 1982, Polak, 1987, Senna, 1994, Noland and Small, 1995, Polak, 1996). Also, a thorough introduction to the topic is (Bates et al., 2001). In summary, the theoretical foundation of the *value of travel time reliability* rests on two frameworks: Centrality-Dispersion (or Mean-Variance) proposed by (Jackson and Jucker, 1982); and Scheduling delays under uncertainty proposed by (Small, 1982) and (Noland and Small, 1995). The first is based on the idea that the travel time unreliability (or variability) is concentrated in a statistical measure of the dispersion of the travel time distribution. The second assumes that travelers have a specified time of arrival, and any *expected* late arrivals or *expected* early arrivals incurs disutilities. These disutilities are asymmetric in contrast to the Centrality-Dispersion framework that assumes all disutilities (due to unreliability) are weighted equally. It should be noted that *expected* refers to the first statistical moment of schedule delays due to late arrivals or early arrivals over the travel time distribution.

The dominant method for the estimation of these values (i.e. VOT and VOR) is discrete choice analysis typically within the Random Utility framework (Ben-Akiva and Lerman, 1985, Train, 2009, Ortuzar and Willumsen, 2011). Generally, the data sources are stated preference experiments, and revealed preference observations. The stated preference experiments present choice scenarios with a variety of presentations (especially in the case of the *value of travel time reliability*) and abstraction (based on real options vs. nondescript options) to travelers. The revealed preference observations refer to actual choices done by travelers in the market (e.g. decisions in the current state of the transportation system of the travelers). Both may also be combined in discrete choice analysis. In revealed preference observations, travel times are experienced by the subjects, and they estimate the travel time through their own cognitive mechanism of perception. This mechanism may be influenced by external sources (e.g. travel information). In essence, there is a mismatch between travel time as reported by a traveler (*subjective travel time distribution*) and travel time as measured from a device (e.g. loop detector; *objective travel time distribution*). It is reasonable that the relationship between subjective travel times and objective travel times may expressed mathematically as: $T_s = T_o + \xi$. T_s is a random variable associated with the probability density given by the *subjective travel time distribution.* T_o is a random variable associated with the probability density given by the *objective travel time distribution.* The variable ξ is the random *perception error* also associated with its own probability density. Thus, it is clear that may overestimate or underestimate the measured travel times, and this is likely to have influence over their valuation of travel time unless $\mathrm{E}(\xi) = 0$, and $\mathrm{Var}(\xi) \approx 0$. In other words, travelers are "optimizing" (i.e. executing decisions on travel choices) according to their own divergent views of the *objective travel time distribution* (Bates et al., 2001, Carrion and Levinson, 2012c).

The primary objective of this study is a systematic comparison of estimated reliability ratios using self-reported travel times from surveys, and measured travel times by Global Positioning System (GPS) devices. The self-reported travel times represent the travelers' *perceived* travel times or the travelers' *subjective* travel times. The measured travel times represent the travelers' *actual* travel time or the travelers' *objective* travel times. The secondary objective is to calculate confidence intervals for the estimated reliability ratios, and observe whether there are overlaps between the confidence intervals of the reliability ratios from self-reported travel times, and reliability ratios from measured travel times. The objectives are accomplished by analyzing data collected of travelers' self-reported travel times, and travelers' measured travel times by GPS devices from a previous research effort (Zhu, 2010, Carrion and Levinson, 2012a). This data is used to estimate two sets of random utility models: systematic utilities with *subjective travel times* (i.e. self-reported travel times); and systematic utilities with *objective travel times* (i.e. GPS measured travel times). Furthermore, the data set consists of the same subjects, and the self-reported and measured travel times for the same trips. Only direct commute trips (from home to work, and from work to home) are considered. The choice dimension is based on the *hierarchy* of the bridges across the Mississippi River in the Minneapolis-St. Paul region. The term *hierarchy* refers whether travelers choose a bridge that belongs to the Dwight D. Eisenhower National System of Interstate and Defense Highways of the United States of America.

5.2 Econometric models

In this study, the data set is analyzed through random utility models (Train, 2009, Ben-Akiva and Lerman, 1985, Ortuzar and Willumsen, 2011). The data set is composed of two observations per subject. There are 39 distinct subjects, and thus 78 observations (see section 3.1, and section 3.6). Each two observation per subject represent the consolidation of all the home to work trips, and all the work to home trips of a subject. The set of home to work trips per subject, and the set of work to home trips per subject allow to obtain travel time distributions for these same trips (see section 3.4) from GPS (*objective travel time distribution*) and survey data (*subjective travel time distribution*) per subject. Thus, centrality and dispersion measures are calculated on the travel time distributions, and are included as attributes in the systematic utilities of the random utility models. The choice dimension is based on the *hierarchy* of the bridges across the Mississippi River in the Minneapolis-St. Paul region (see figure 3.1). The term *hierarchy* refers whether travelers choose a bridge (from those listed in figure 3.1) that belongs to the Dwight D. Eisenhower National System of Interstate and Defense Highways of the United States of America. It should be noted that only 39 subjects had enough observations on Interstate bridges, and non-Interstate bridges. Thus, the set of home to work trips and the set of work to home trips are further disaggregated to two alternatives (or choices). The first alternative represents the *most used* bridge that belongs to the Interstate category, and the second alternative represents the *most used* bridge that belongs to the non-Interstate category. The term *most used* refers to the bridge with the highest number of commute trips. The bridge with the smallest travel time by centrality measure and dispersion measure is selected in the case that two or more bridges have the same number of commute trips. Furthermore, an alternative is considered chosen by a subject according to whether the number of trips on the alternative is strictly higher compared to the other alternative.

5.2.1 Random utility models

The random utility models considered in this study can be formulated as binomial logits (Train, 2009). Assume that the utility function a decision-maker n in the set of decision-makers $\mathcal{N}$ associates with alternative j in the set of choices $\mathcal{C}_n$ (for this study $\mathcal{C}_n$ only have two alternatives) is given by:

$$U_{nj} = V_{nj} + \epsilon_{nj} \tag{5.1}$$

For this case of binomial logit model, the functional form is given by equation 5.1. The first term (V_{nj}) is the systematic utility, and the second (ϵ_{nj}) is a random vector identically and independently distributed (i.i.d.) over alternatives and decision-makers following a extreme value type 1 (or Gumbel) distribution with 0 location, and scale set to 1. For this study, the systematic utility is linear in parameters; $V_{nj} = \beta^T x_{nj}$, where β is the coefficient vector, and x_{nj} are the vectors of explanatory variables in the regressors matrix.

5.2.2 Systematic Utility for the models

The additive linear in parameters systematic utility for the alternatives for all models is:

$$V_{nj} = f(T, V, S, D, A; \beta) \tag{5.2}$$

where

- T: Centrality measure of travel time
- V: Dispersion measure of travel time
- S: Socio-demographic
- D: Type of work trip
- A: Alternative specific constants (ASC)

Centrality measure of travel time

The centrality measures are calculated for the travel time distributions for the set of home to work trips, and work to home trips for each alternative per subject as described in section 5.2. For this study, the mean, and the median are considered as centrality measures. These variables are alternative specific. The variables are measured in minutes.

Dispersion measure of travel time

The dispersion measures are calculated for the travel time distributions for the set of home to work trips, and work to home trips for each alternative per subject as described in section 5.2. For this study, the standard deviation (a typical measure in the Centrality-Dispersion framework), and the difference between the 90th percentile and the median (DMP90) are considered as dispersion measures. These variables are alternative specific. The variables are measured in minutes.

Socio-demographic

These are extracted from the socio-demographic questions in the web-based surveys.

- Gender (1 = Male; 0 = Female).
- Income. Four categories: $(\$0, \$49,999]$, $(\$50,000, \$74,999]$, $(\$75,000, \$99,999]$, and $(\$100,000, \infty+)]$. The first category is the base case. (2008 US dollars).

Type of work trip

It is a binary variable indicating whether the trip originates from home (1 = from home to work) or from work (0 = from work to home).

Alternative specific constants

For these binomial logits, the alternative specific constant of the Interstate alternative is set to 0.

5.2.3 Statistical hypothesis testing

There are two hypothesis tests that are considered for the random utility models in this study. For the nested models, the *Wald tests* are used as they only depend on the covariance matrix of the unrestricted models, and do not require estimation of the restricted models. These tests are asymptotically equivalent to the *likelihood ratio tests*. For the nonnested models, the *Akaike information criterion* (AIC), and *Bayesian information criterion* (BIC) are used in order to compare the statistical fit of the binomial logits with travel times from survey data to the binomial logits with travel times from GPS data. Furthermore, the confidence intervals for the *reliability ratio* of the models are calculated using the *Delta method* (Cramer, 1986, Johnston and DiNardo, 1997, Pawitan, 2001, Greene, 2012).

5.2.4 Estimation

The estimation of binomial logits are straightforward, and it is done by maximizing the loglikelihood, which is of closed form. The details are standard and are found in (Ben-Akiva and Lerman, 1985, Train, 2009, Ortuzar and Willumsen, 2011). The models are estimated using STATA (Cameron and Trivedi, 2010).

The likelihood for these binomial logit models is given by:

$$L(\beta) = \prod_{\forall n \in \mathcal{N}} \prod_{\forall j \in \mathcal{C}_n} \left(\frac{e^{V_{nj}(\beta)}}{\sum_{j \in \mathcal{C}_n} e^{V_{nj}(\beta)}} \right)^{\delta_{nj}} \tag{5.3}$$

Where the δ_{nj} variable is one for the chosen j alternative of the n decision-maker, and zero otherwise.

5.3 Discussion and results

Table 5.1 presents the estimates of the random utility models (binomial logits) along with the reliability ratios, and goodness of fit statistics. There are four types of models estimated according to distinct centrality and dispersion measures of travel time: Mean/Standard Deviation (SD); Mean/Difference between 90th percentile and median (DMP90); Median/SD; and Median/DMP90. The four types of models are estimated with self reported travel times from surveys, and with measured travel times from GPS devices. The results indicate that the estimates of the centrality measures and dispersion measures of travel times are negative, and highly statistically significant across all models. The goodness of fit statistics indicate that the models with self reported travel times from surveys fit the data better in contrast to models with measured travel times from GPS devices. Both the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC) favor the models with self-reported travel times over the models with measured travel times. Thus, the models estimated with self reported travel times are preferred by statistical basis, and more specifically the Mean/SD model.

The reliability ratios in the models with self reported travel times are higher than 1, except for the Mean/DMP90. The 95% confidence intervals of these models indicate that values greater than 1 are more plausible. In contrast, the reliability ratios in the models with measured travel times are less than 1. The 95% confidence intervals of these models indicate that values less than 1 are more plausible. In addition, there are few overlaps of the 95% confidence intervals for the same models with self reported travel times vs. measured travel times. This is an important finding as it gives different results with regards to the subjects valuing of travel time savings, and travel time reliability.

The reliability ratio is defined as the marginal rate of substitution between the travel time variability, and the expected travel time. Thus, the centrality-dispersion models with the self reported travel times indicate that the subjects are valuing higher the travel time variability over the expected travel time, except for the Mean/DMP90 model. In contrast, the centrality-dispersion models with measured travel times indicate that the subjects are valuing higher the expected travel time over the travel time variability. Therefore, questions arise about which of the travel times (self reported or measured) should be trusted. This leads back to the previous discussion about perception error. It is known that perception error is a factor that distorts the travelers' interpretation of the actual travel times. It is more than likely that the travelers execute their travel decisions based on their perceived travel times, and not the actual travel times. This perception also is linked to the valuation of travel time, and thus the reliability ratios may be inflated or deflated depending on the level of distortion or magnitude of the perception error.

Lastly, the socio-demographic (e.g. income and gender), and type of work trip variables were not found statistically significant. Thus, the subjects were more influenced by the travel time measures in their choices. This result agrees with the findings in Carrion and Levinson (2012a) with the same data source, albeit not the exact same data set.

5.4 Conclusion

This study presents novel results that are starting to scratch the surface of the influence of perception on the valuation of travel time. At the moment, there are few intersections of two main research areas in the transportation literature: travelers' perception of travel time; and travelers' valuation of travel time with a greater emphasis on the valuation of travel time reliability. There are already many studies identifying that subjects' perception of travel times has been found to be a significant factor in studies. Travelers overestimate or underestimate the actual travel times they experience. Therefore, it is likely that revealed preference studies may be underestimating or overestimating the value of travel time savings, and value of travel time reliability as the objective travel time distributions (measured from devices) differ from the subjective travel time distributions (self reported by travelers).

In this study, the influence of commuters' perception error is investigated by estimating random utility models (i.e. econometric models) on data collected of commuters recruited from a previous research study in the Minneapolis-St. Paul region (Zhu, 2010, Carrion and Levinson, 2012a). This data (surveys, and Global Positioning System [GPS] points) consists of work trips (from home to work, and from work to home) of subjects . For these work trips, the subjects' self-reported travel times, and the subjects' travel times measured by GPS devices were collected. The results indicate that the subjects value travel time reliability more than travel time savings (i.e. reliability ratios greater than 1) in the econometric models with self-reported travel times. In contrast, subjects value travel time savings more than travel time reliability (i.e. reliability ratios smaller than 1) in the econometric models with travel times as measured by GPS devices. Furthermore, the models with self reported travel times are found to statistically fit the data better in comparison to the models with measured travel times.

Finally, these initial results are actually a harbor for the departure of new studies trying to "unpack" the factors governing the perception error of the travelers, and also for studies focusing on the influence of external sources of information (e.g. travel information) on the magnitude of the values of travel time savings, and values of travel time reliability.

Table 5.1: Random Utility Models

Variables	Survey (Mean/SD)	Survey (Mean/DMP90)	Survey (Median/SD)	Survey (Median/DMP90)	GPS (Mean/SD)	GPS (Mean/DMP90)	GPS (Median/SD)	GPS (Median/DMP90)
	Estimates (T-Stats)	Estimates (T-Stats)	Estimates (T-Stats)	Estimates (T-Stats)	Estimates (T-Stats)	Estimates (T-Stats)	Estimates (T-Stats)	Estimates (T-Stats)
Centrality - Travel time - [Interstate/non-Interstate]	-1.19 (-2.89) ***	-0.46 (-3.72) ***	-1.20 (-2.90) ***	-0.70 (-3.56) ***	-0.33 (-3.88) ***	-0.31 (-3.86) ***	-0.19 (-3.39) ***	-0.17 (-3.31) ***
Dispersion - Travel time - [Interstate/non-Interstate]	-1.54 (-2.89) ***	-0.38 (-3.41) ***	-1.93 (-3.05) ***	-0.61 (-3.71) ***	-0.14 (-3.01) ***	-0.12 (-2.75) ***	-0.15 (-3.41) ***	-0.11 (-2.91) ***
Gender - [non-Interstate] 1 = Male; 0 = Female	1.25 (0.85)	1.54 (0.103)	2.02 (1.46)	1.84 (1.69) *	-0.54 (-0.75)	0.25 (0.52)	0.09 (0.14)	0.55 (0.94)
Income - [non-Interstate] ($50,000, $74,999] 1 = In; 0 = Out	1.32 (0.76)	0.400 (0.36)	0.66 (0.42)	0.05 (0.04)	0.68 (0.71)	-0.13 (-0.14)	-0.39 (0.44)	-0.63 (-0.76)
Income - [non-Interstate] ($75,000, $99,999] 1 = In; 0 = Out	1.52 (0.86)	-0.042 (-0.04)	0.66 (0.42)	-0.29 (-0.26)	0.14 (0.14)	-0.80 (-.78)	-0.27 (-0.31)	-0.79 (-0.92)
Income - [non-Interstate] ($100,000, ∞+)] 1 = In; 0 = Out	-1.75 (-0.77)	-1.03 (-0.89)	-3.73 (-1.65) *	-1.66 (-1.16)	1.07 (1.17)	0.25 (0.30)	0.08 (0.10)	-0.32 (-0.43)
Type of work trip - [non-Interstate] 1 = from home to work; 0 = from work to home.	-0.19 (-0.14)	0.49 (0.62)	0.29 (0.80)	0.75 (0.80)	-0.29 (-1.37)	0.35 (0.52)	-0.16 (-0.28)	0.31
Alternative Specific Constant - [non-Interstate]	-2.07 (-1.13)	-2.02 (-1.86) *	-2.04 (-1.48)	-2.35 (-1.85) *	-1.03 (-1.37)	-1.20 (-1.66) *	-0.22 (-0.31)	-0.60 (-0.90)
Reliability Ratio RR 95% Confidence Interval	1.30 [1.08, 1.51]	0.83 [0.59, 1.08]	1.60 [1.32, 1.89]	1.14 [0.95, 1.35]	0.42 [0.18, 0.69]	0.39 [0.16, 0.61]	0.77 [0.37, 1.16]	0.65 [0.28, 1.02]
Intercept Log-Likelihood $ll_{A\hat{S}C}$	-51.472515	-51.472515	-51.472515	-51.472515	-51.472515	-51.472515	-51.472515	-51.472515
Final Log-Likelihood $ll_{\hat{\beta}}$	-9.4065462	-22.884538	-10.486297	-18.352769	-30.111183	-31.944981	-36.652578	-39.258894
Likelihood ratio index ρ^2	0.81725109	0.55540276	0.79627386	0.64344525	0.41500463	0.37937789	0.28791942	0.23728433
Akaike Information Criterion AIC	34.81309	61.76908	36.97259	52.70554	76.22237	79.88996	89.30516	94.51779
Bayesian Information Criterion BIC	59.21194	86.16792	61.37144	77.10439	100.6212	104.2888	113.704	118.9166
Number of observations	78	78	78	78	78	78	78	78
Number of subjects	39	39	39	39	39	39	39	39

* is 10% significance level, ** is 5% significance level, *** is 1% significance level

See the section 5.2 for details on the econometric models.

Chapter 6

Route choice dynamics after a link restoration

6.1 Introduction

The most basic wisdom of travel behavior is that travelers adapt to their circumstances according to their own knowledge inside the road network. This knowledge is the outcome of human-environment interaction related to the act of traveling. Travelers refine their movements in their surroundings through spatial, and temporal information acquisition. This information is connected to two guiding processes: navigation, and wayfinding. Navigation describes the actions required for unobstructed movement by linking locations of places, and trajectories between places. Wayfinding is about selecting routes connecting an origin-destination pair of interest to the traveler. In short, travelers learn about some of the places connected by the transportation system. Travelers learn about some of the distinct alternatives (i.e. mode, route) to navigate in the transportation system. Travelers learn about some of the states (e.g. peak hour congestion) of the transportation system at specific times during the day, week, month, and in some cases even year. Furthermore, travelers' may exercise any combination of several possible responses available to them that vary according to timing. In the short term, travelers potential responses include: rescheduling trips to earlier or later times; switching routes; canceling trips; and others. In the long term, travelers potential responses include: auto ownership; finding alternative location of activities;

moving to a new residential and/or work location; and others. It is important to realize that travelers choose a proper bundle of potential responses depending on the characteristics of the travelers themselves, and of the physical environment. The characteristics of the travelers consist of objective socio-demographic elements (age, gender, income,...) and subjective elements (preferences, perception, experiences...). The physical environment is characterized by the built-up surroundings: the transport network infrastructure connecting the set of places; and the set of places including but not limited to: residential locations, commercial locations, parks, and others. The characteristics of the built-up surroundings (transportation network infrastructure and/or the set of places) are numerous, but only a subset have been found relevant to travel behavior. This subset includes: travel time; travel time variability/reliability; travel cost; travel distance; aesthetics of scenery; network structure; traffic information; and others. The interrelations between the characteristics of the travelers, and of the physical environment define the behavior of the travelers. In essence, the travelers perceive the characteristics of the physical environment, and the travelers extract the relevant information according to their own criteria. This information is processed also according to their own criteria, and a bundle of possible responses is chosen. Lastly, this selection process is dynamic; it receives feedback (e.g. past experience) from the travelers' previous decisions (Ben-Akiva et al., 1984, Bovy and Stern, 1990, Golledge, 1992, Golledge and Stimson, 1997, Golledge, 1999).

This study focuses on uncovering the dynamics of travel behavior, more specifically, the dynamics of bridge choice behavior after a large-scale disruption. It investigates the day-to-day behavior of commuters after the opening of the replacement bridge for the previously collapsed I-35W bridge in the Minneapolis-St. Paul metropolitan region. The original I-35W bridge collapsed on August 1st 2007, and the replacement bridge opened to the public on September 18th 2008. The primary objective of this study is to identify the factors that influence the day-to-day subjects' decision to stay or abandon their current chosen bridge, and in addition to determine the possible relationships between these factors. For this purpose, this study analyzes data collected of commuters recruited from a previous research effort (Carrion and Levinson, 2012a). This data consists of Global Positioning

System (GPS) points, and web-based surveys. This data was collected before and after the replacement bridge opened. The GPS data of the subjects contains geographical points between the last weeks of August 2008, and the first weeks of December 2008.

Carrion and Levinson (2012a) only focused on analyzing the traffic patterns of the subjects on the road network, and identifying the possible reasons influencing the subjects' preferences towards the new I-35W bridge vs. other bridge alternatives. The preferences of a subject are ascertained by the proportion of trips on the I-35W bridge out of the total trips of the subject vs. the proportion of trips on all the other bridge alternatives out of the total trips of the subject. Furthermore, the model, besides socio-demographic variables, also incorporated statistical measures (e.g. mean, standard deviation) of day-to-day travel time distributions. The travel time distributions are obtained by aggregating travel times of different days for the distribution of the I-35W bridge, and for the distribution for all the other alternatives. Thus, Carrion and Levinson (2012a) assumed implicitly that subjects at the aggregate level settle for a particular bridge choice. However, this assumption neglects that subjects may be exercising their decisions at a day to day level. This study extends Carrion and Levinson (2012a) by considering explicitly the day-to-day behavior of travelers, and by also considering the previously excluded subjects that are transitioning between bridge alternatives not including the I-35W bridge. This is accomplished by specifying and estimating a *duration model* (i.e. a hazard model) on data of the subjects' morning commute.

6.2 Limitations and discussion of current research

In the transportation research literature, empirical models that incorporate travel time reliability are *static models*, and *Random Utility models* (Ben-Akiva and Lerman, 1985, Train, 2009, Ortuzar and Willumsen, 2011). The models are *static*, because they only consider trade-offs of statistical measures on travel time distributions of a set of alternatives (e.g. routes, bridges). The travel time distributions may be hypothetical (stated preference data) or day-to-day travel times (revealed preference data). In stated preference studies, the travel time distribution of each alternative is generated by the researchers. In addition, the researchers are familiar with the value of different statistical measures (e.g. mean, median), based on the previous theoretical frameworks (see section 2.2), of the travel time distribution of each alternative. Thus, subjects simply observe a set of attributes of each alternative that are abstractions of the concept of travel time reliability/variability, and of expected travel time. These abstractions depend on different presentations (e.g. numbers and/or visual aid), and in cases may include the travel time distributions of the alternatives such as in the case of histograms in the stated choice questions. On the other hand, revealed preference studies produce travel time distributions for an alternative by aggregating travel times of different days of the alternative. Statistical measures are computed on these day-to-day travel times, and it is assumed that these measures represent the experience of the subjects with regards to the travel time reliability/variability, and the expected travel time for each alternative. In essence, the dynamic behavior of the travelers have been neglected to favor an assumption that subjects, at the aggregate level, settle for a particular alternative. However, this assumption neglects that subjects may be exercising their decisions at a day to day level. In addition, the subjects may only consider a subset of the day to day travel times for each alternative, and that this subset of day to day travel times is updated every few days. In other words, subjects may consider adding to the set day-to-day travel times according to their perception, and the subjects may discard day-to-day travel times according to their limited memory. Furthermore, the models are *Random Utility models*, and thus they are bound by the assumption that travelers as utility-maximizers know the expected travel time, and the travel time variability of each alternative. Travelers are *informed* with regards to the attributes related to travel time (i.e. means, standard deviation) for each alternative. This is a fundamental assumption in *Random Utility models* based on the theory of rational choice presented in (Domencich and McFadden, 1975). It is possible to include only the travel times known for certain alternatives, but this implicitly assumes that the travel times of the alternatives, that excludes them, are of the highly unlikely value of zero. This is also a questionable assumption. In fact, questions arise about whether a subject truly knows of the travel time of his other alternatives, or a subject only knows the travel time obtained from past accumulated experience. In addition, questions arise about the modus operandi of a subject to choose whether to stay or abandon the current chosen alternative based on its day to day travel times. In light of this discussion, researchers must ask whether stated preference studies are presenting a decision-making situation that is *realistic*, and whether revealed preference studies with travel time distributions obtained by aggregating day to day travel times are *simplistic*. Lastly, it may be argued that the few studies with *dynamic Random Utility* models are not able to overcome discussed issues satisfactorily. Stated preference studies with multiple choice situations are considered more likely to be cases of unobserved heterogeneity rather than dynamics. Revealed preference studies are able to include the dynamics by calculating the statistical measures on day-to-day travel time distributions that are updated during each choice situation of a decision-maker. This update is accomplished by including the new travel times for each choice situation. Unfortunately, it requires significant data collection by the researchers of trips of subjects on each alternative, and the assumption that subjects are *informed*, and always remember all past travel times are still present. It should be noted that these concerns are similar to those put forward by Simon (1997) in the theory of *bounded rationality*.

6.3 Econometric models

In this study, the data set is analyzed through *duration analysis* (Kalbfleisch and Prentice, 2002, Cameron and Trivedi, 2005, Wooldridge, 2010). The dependent variable is the *single-spell duration* per subject. This duration is defined as the date and time elapsed from September 18th 5:00 AM until the date and time a subject *consistently* leaves his *current bridge choice* for any other of those in figure 3.1 (transition is observed), or the date and time until the GPS device is retrieved from the subject, and the subject has not left his *current bridge choice* (transition is not observed). The term *current bridge choice* refers to the subject's bridge choice at or after September 18th 5:00 AM. The term *consistently* refers to a subject's transition from his current bridge choice to another bridge at least two times consecutively. The term *single-spell duration* refers to modeling only one single transition from the *current bridge choice* to any other of those in figure 3.1. This single transition is only the first transition observed in the subjects' GPS data. In the data set, there are only 65 subjects (see section 3.6), and only 65 single-spell transitions (one single-spell observation per subject). There are 835 observations, and on average 12.84 observations per subject. In addition, subjects are observed on average for 32.06 days, and 34 subjects are observed to *consistently* transition from their *current bridge choice* to another bridge of those in figure 3.1. Also, the order of the events (i.e. different observed days and times of subjects' trips) of the *single-spell duration* per subject is known to the minute (year, month, day, hour, and minute).

6.3.1 Duration models: Basic concepts

The *single-spell duration* of a subject n (n refers to a generic subject from the sample) is assumed as a nonnegative random variable (T_n). This random variable T_n has a *cumulative distribution function* (or probability distribution) denoted $F_n(t)$, and has a *probability density function* denoted $f_n(t)$ or $\frac{dF_n(t)}{dt}$. Therefore, the probability that the *single-spell duration* of a subject n is less than a specific date and time t is given by

$$F_n(t) = Prob[T \leq t] \tag{6.1}$$

$$F_n(t) = \int_0^t f_n(u)du \tag{6.2}$$

Another important concept is the probability that the *single-spell duration* of a subject n equals or exceed a specific date and time t is known as the *survivor function* ($S_n(t)$). Thus, the *survivor function* is a monotone, and non increasing function of time that reports the probability of a subject that still has not transition to another bridge beyond a specific date and time t. This is defined formally as

$$S_n(t) = Prob[T \geq t] \tag{6.3}$$

$$S_n(t) = 1 - F_n(t) = \int_t^\infty f_n(u)du \tag{6.4}$$

The *hazard function* of subject n ($h_n(t)$) is another important concept that refers to the instantaneous rate of transition. It is the limit of the probability that a transition occurs in a specific interval conditioned on the probability that the subject still has not transition at least until the start of the interval. For the *hazard function*, a value of zero means no transition at the interval, and a high value close to ∞ means a certainty of transition. *Hazard functions* may be constant, increasing or decreasing as functions of the date and time of transition. This is defined formally as

$$h_n(t) = \lim_{\Delta t \to 0} \frac{Prob[t + \Delta t > T > t | T > t]}{\Delta t} = \frac{f_n(t)}{S_n(t)} \tag{6.5}$$

Lastly, the *cumulative hazard function* is a measure of the total hazard that has accumulated up to a specific date and time t. This is defined formally as

$$H_n(t) = \int_0^t h_n(u) du = \int_0^t \frac{f_n(u)}{S_n(u)} du = -\ln(S_n(t)) \tag{6.6}$$

The general assumption of duration models, similar to other econometric models, is that the subjects' single-spell durations (T_n) are independent and identically distributed with probability density function $f(t|\beta, x)$ (i.e. the n is thus unnecessary) conditional on the parameters and covariates. In addition, Maximum Likelihood is the generally used estimation method for duration models (Cramer, 1986, Pawitan, 2001).

6.3.2 Cox Proportional Hazard model

Duration models, similar to other econometric models, may have nonparametric, semi-parametric, and/or parametric models for the dependent variable. The Cox Proportional Hazard (hereafter referred as Cox PH) model is a semi-parametric model that assumes the hazard function (conditional on the parameters and covariates; $h(t|\beta, x)$) is factored into two separate functions (proportional hazards assumption).

$$h(t|\beta, x) = h_0(t)\phi(x, \beta) \tag{6.7}$$

$h_0(t)$ is known as the *baseline hazard*, and $\phi(x, \beta)$ is known as the *relative hazard*. The baseline hazard is a function of the date, and time of transition, and its functional form is left unspecified. The shape of the baseline hazard function depends on the data, and it is not assumed as in fully parametric models. The baseline hazard function is estimated using a nonparametric product limit estimator similar to the Kaplan-Meier estimator (Cameron and Trivedi, 2005, Kalbfleisch and Prentice, 2002). This is further explained in section 6.3.5. The relative hazard is a nonnegative function (i.e. hazard rates cannot be negative) that is fully specified by the researcher. The common functional form of $\phi(x, \beta)$ is the exponential ($\phi(x, \beta) = e^{\beta^T x}$; β is a vector of coefficients, and x are the vectors of covariates in the regressors matrix). The relative hazard is estimated by maximizing a Partial Likelihood function. This is further explained in section 6.3.5. For this study, the hazard rate is

$$h(t|\beta, x) = h_0(t)e^{\beta^T x} \tag{6.8}$$

The basic assumption of the Cox PH's hazard function is that all subjects have the same baseline hazard function ($h_0(t)$), and the relative hazard ($\phi(x, \beta)$) has a multiplicative effect on the baseline hazard depending on the values of the covariates. In other words, one subject's hazard function is a multiplicative version of another subject's hazard function. This is the proportional hazards assumption, and it is statistically tested for the data set of this study. This is further explained in section 6.3.4 along with other statistical tests. Furthermore, the Cox PH model may be adjusted to consider: tied observations; and censorship of single-spell durations. The tied observations refer to the lack of information with regards to the order of the subjects that transition at the same date, and time in the data set. The order matters as it is required for the estimation of the Cox PH model. In this study, there are no ties as the order of the events (i.e. different observed days and times of subjects' trips) of the *single-spell duration* per subject is known to the minute (year, month, day, hour, and minute). The censorship (or more precisely right-censorship) of single-spell durations refer to the researchers not observing the transition of the subjects. An observed transition is defined as a subject *consistently* leaving his *current bridge choice* for any other of those in figure 3.1. An unobserved transition is a subject not leaving his *current bridge choice* before the GPS device is retrieved from the subject. In other words, the study ends before a transition is observed. Readers should remember that the term *current bridge choice* refers to the subject's bridge choice at or after September 18th 5:00 AM. In this study, there are 34 subjects (out of a total of 65 subjects) are observed to *consistently* transition (i.e. at least two times consecutively) from their *current bridge choice* to another bridge of those in figure 3.1. In contrast, there are 31 subjects with unobserved transition. It is assumed that the censorship mechanism is independent from the single-spell duration of the subjects. This is a fair assumption given that the time that a GPS device is retrieved from a subject did not depend on whether a subject changed bridge choices or not, but rather on the fixed time duration of the study (i.e. 8 or 13 weeks depending on the data collection effort; see section 3.1).

6.3.3 Relative hazard: covariates

The selection of the covariates in the relative hazard function is based on two major groups: characteristics of the subjects; and travel time measures. The characteristic of the subjects are obtained from web-based survey data, and GPS data. The travel time measures are calculated on travel times obtained from GPS data. Readers may refer to section 3.4 for details. In addition, there are two types of covariates: time-invariant (y); and time-dependent ($w(t)$). Thus, the relative hazard is given by

$$\phi(y, w(t); \beta_y, \beta_w) = e^{\beta_y^T y + \beta_w^T w(t)} \tag{6.9}$$

β_y, and β_w are vector of coefficients to be estimated. y is a vector of time-invariant covariates (only change across subjects, and not across subjects' day-to-day morning commute trips). $w(t)$ is a vector of time-dependent covariates (the values vary across subjects' day-to-day morning commute trips).

The time-invariant covariates (y) are:

- y_1: Past bridge diversity.
- y_2: Gender.

- y_3: Income.
- y_4: Ratio of bridge distances (current bridge choice to previous bridge choice).
- y_5: Fear of driving on the I-35W bridge and other bridges in the vicinity.

Past bridge diversity (y_1)

The number of distinct alternatives (bridges) a subject used for his morning commute trip before September 18th 5:00 AM. This covariate is an indication of a subject's knowledge of alternative bridges before traveling each day from September 18th 5:00 AM until the subject decided to change to another bridge alternative.

Gender (y_2)

It is a binary variable; 1 = Male; 0 = Female.

Income (y_3)

It is a set of three binary variables: Low income ($\$0, \$49,999$], Medium income ($\$50,000, \$99,999$]. and High income ($\$100,000, \infty+$). The first category is the base case (2008 US dollars).

Ratio of bridge distances (current bridge choice to previous bridge choice; y_4)

It is the ratio of Euclidean distances. The numerator is the Euclidean distance from a subject's home location to the centroid of the *current bridge choice*, and from this centroid to the subject's work location. The denominator is the euclidean distance from a subject's home location to the centroid of the *previous bridge choice*, and from this centroid to the subject's work location. The term *current bridge choice* refers to the subject's bridge choice at or after September 18th 5:00 AM. The term *previous bridge choice* refers to the subject's bridge choice with the highest number of trips before September 18th 5:00 AM.

This covariate measures a relative impedance of distance between the current bridge, and an alternative bridge highly preferred in the past by the subject.

Fear of driving on the I-35W bridge and other bridges in the vicinity (y_5)

It is a binary variable. This variable identifies the subjects that admitted they avoid bridges (including the I-35W bridge, Washington Ave bridge, and 10th Street bridge), because of fear of bridge collapse or any other reason in the web-based surveys.

The time-dependent covariates ($w(t)$) are the following travel time measures:

- $w_1(t)$: Fixed thresholds
- $w_2(t)$: Moving thresholds
- $w_3(t)$: Travel times from the previous most used bridge

Travel time measures

It is challenging to come up with measures that capture the relevant aspects of travel time considered by travelers. In a previous discussion (see section 6.2), the authors identified several key difficulties with the current models incorporating travel time measures representing both expectation, and variability of travel time. In summary, travelers *perceive* the travel times at the end of each of their trips, and travelers *recall* travel times from past trips. Thus, two important concerns are: the selection of day to day travel time by travelers due to their perception; and the ability of the travelers to recall the previously selected day to day travel times.

In this study, the authors hypothesize that the subjects only consider a subset of their travel times, and that the subjects only recall every few days a portion of the travel times in this subset. In addition, the travel times in this subset are continuously updated according to thresholds set by the subjects. Figure 6.1a presents the day to day travel times of a subject in the sample. A kernel weighted local polynomial smoother (epanechnikov; 5.29 bandwidth) with its 95% confidence interval is fitted to the day to day travel times to further elucidate the trend across days. The smoother indicates that the travel times are constantly in flux until they finally increase significantly. Therefore, it must be asked how do travelers react to the trend and volatility of the observed travel time time series. It is reasonable given the *perception* of travelers that only certain travel times have weight in influencing the choices of travelers. In other words, travel times across days that are closely similar (e.g. 20 minutes and 24 minutes) may not be as noticeable to the traveler in comparison to travel times across days that are quite dissimilar (e.g. 20 minutes and 35 minutes). Thus, it is assumed that the subjects consider travel times above a certain number of standard deviations from the mean, and also travel times below a certain number of standard deviations from the mean. The standard deviation is calculated on the travel times of the travelers' trips, and the mean is calculated on the travel times of the travelers' trips. Two possible rules for thresholds are: fixed, and moving. The *fixed thresholds* (see figure 6.1b) assume travelers have an rigid expectation of the travel time, and a rigid magnitude for the travel time variability. Travel times within margins (e.g. mean plus one standard deviation, and mean minus one standard deviation) are acceptable and travel times above are not desirable, and travel times below may also not be preferred. In addition, the margins may be asymmetric indicating that travelers may be more forgiving of travel times below the mean, but not so much of travel times above the mean. The *moving thresholds* (see figure 6.1c) assumes travelers' expectation of travel time, and travelers' variability of travel time are continuously updated according to the travelers' travel times of past trips. The mean is a moving mean that may only consider the travel time of two or more previous trips, and the standard deviation is a moving standard deviation that may also only consider the travel time of two or more previous trips. The shorter the number of previous trips considered the closer is the moving mean to the actual travel times. The margins are continuously updated, and thus travel times, that may not be acceptable (i.e. within the thresholds) in past trips, are acceptable in future trips. Furthermore, travelers, besides having fixed or moving thresholds, may also have a *fixed* or *moving* set of travel times that they can remember or recall. Travelers may discard (or forget) the travel times, and the frequency of past trips beyond their thresholds, and within their thresholds. This set of travel times is likely to be updated across days, and thus effectively travelers may only remember travel times and frequencies of past trips up to a specific number of days.

Lastly, the influence of the travelers' past trips on their *previous most traveled bridge choice* is considered. It is plausible that travelers at the beginning of their trips in their *current bridge choice* may contrast the current travel experience on the current bridge with the past travel experience on the past bridge. In addition, the travelers may eventually forget about their past travel experience on the past bridge. Also, it is plausible that travelers may change to other bridges, because they are familiar with these bridges without necessarily remembering their previous travel experience, or possessing a travel experience on these bridges. It is hypothesized that travelers compare the difference between the median of the travel times of their previous most traveled bridge choice with the each of the travel times experienced by the travelers for each trip.

The proposed travel time measures of this study are: *fixed thresholds* defined by mean, and number of standard deviations from the mean; *moving thresholds* defined by moving means, and a number of moving standard deviations from the moving mean; and travel times from the previous most used bridge.

Fixed thresholds ($w_1(t)$)

The mean, and the standard deviation of the travel times of a subject's set of morning commute trips are calculated. Furthermore, the authors test thresholds that are 0.5, 1, 2, and 3 times from the mean. The authors test both symmetric thresholds (e.g. mean - 0.5 standard deviation, and mean + 0.5 standard deviation), and asymmetric thresholds (e.g. mean - 0.5 standard deviation, and mean + 1 standard deviation). In addition, a moving set of trips is considered. This moving set include trips that are from 2 to 15 days ago from the specific day of travel of a trip. In other words, the set considers exactly a fixed number of trips right before each of current trips a subject undertook. These variables measure the proportion of trips (number of trips divided by total number of trips) that are above (e.g. travel time greater than mean plus one standard deviation), and below (e.g. travel time less than mean plus one standard deviation) the thresholds. Proportion of trips within the thresholds are considered. The authors refer to trips above the thresholds as *late trips*, and to trips below the thresholds as *early trips*. Trips within the thresholds are *regular trips*.

Moving thresholds ($w_2(t)$)

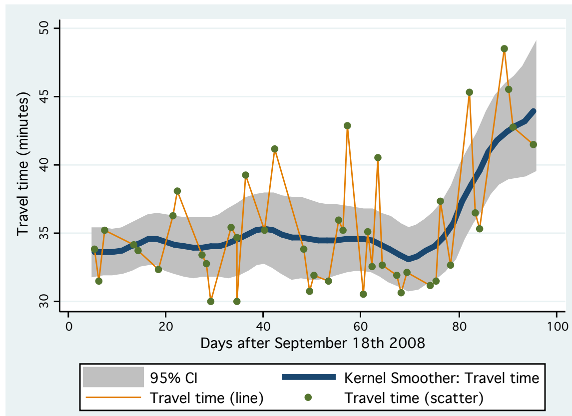
The moving mean, and the moving standard deviation of the travel times with distinct thresholds that are 0.5, 1, 2, and 3 times from the moving mean are considered. The moving mean, and the moving standard deviation are calculated on a moving set that includes trips that are from 2 to 15 days ago from the specific day of travel of a trip. The moving set considers exactly a fixed number of trips right before each of current trips a subject undertook. These variables measure the proportion of trips (number of trips divided by total number of trips) that are above (e.g. travel time greater than moving mean plus one moving standard deviation), and below (e.g. travel time less than moving mean plus one moving standard deviation) the thresholds. Proportion of trips within the thresholds are considered. The authors refer to trips above the thresholds as *late trips*, and to trips below the thresholds as *early trips*. Trips within the thresholds are *regular trips*.

Travel times from the previous most used bridge ($w_3(t)$)

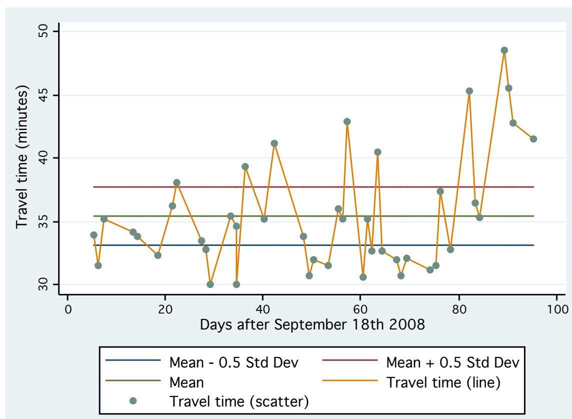
The bridge with the highest number of trips before September 18th 5:00 AM is identified. The median of the travel times of the different days a traveler used the bridge is calculated. The difference between this median, and each of the days travel times is computed. Furthermore, it is assumed that the coefficient of this time-dependent covariate is also time-dependent. Basically, the coefficient for the first week of travel of a subject in the current bridge choice is different from the coefficient after the first week of travel in the current bridge choice.

Figure 6.1: Thresholds-based behavior of travelers

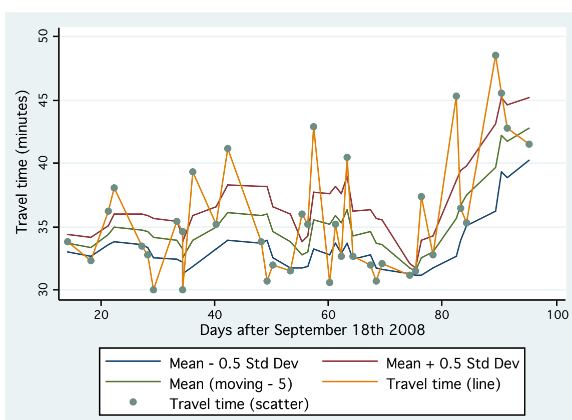
(a) Day to day (GPS) travel times of a subject



(b) Fixed thresholds



(c) Moving thresholds



6.3.4 Statistical hypothesis testing and goodness of fit

There are two hypothesis tests that are considered for the duration models in this study. For the nested models, the *Wald tests* are used as they only depend on the covariance matrix of the unrestricted models, and do not require estimation of the restricted models. These tests are asymptotically equivalent to the *likelihood ratio tests*. For the nonnested models, the *Akaike information criterion* (AIC), and *Bayesian information criterion* (BIC) are used in order to compare the statistical fit of the duration models with the different travel time measures proposed in section 6.3.3 (Cramer, 1986, Johnston and DiNardo, 1997, Pawitan, 2001, Greene, 2012).

The goodness of fit statistics of the models are tested using residual analysis (Cameron and Trivedi, 2005). The Schoenfeld residuals (Schoenfeld, 1982, Grambsch and Therneau, 1994) are computed to test the proportional hazards assumption required for the Cox PH model, and the Deviance residuals (Collett, 2003) are examined to test the model accuracy and to identify outliers.

6.3.5 Estimation

The estimation of the Cox PH model is done in two stages. The first step is the maximization of a Partial Likelihood function to obtain estimates of the coefficients, and the covariance matrix for the coefficients in the relative hazard function. The second is the maximization of a product limit estimator to obtain the hazard rate contributions given the estimates of the relative hazard. This estimator is a nonparametric Likelihood function (Kalbfleisch and Prentice, 2002, Cameron and Trivedi, 2005).

The Partial Likelihood function for this study is

$$L(\beta_y, \beta_w) = \prod_{j=1}^{N} \frac{e^{\beta_y^T y_j + \beta_w^T w_j(t)}}{\sum_{m \in R(t_j)} e^{\beta_y^T y_m + \beta_w^T w_m(t)}} \tag{6.10}$$

The N is the number of *single-spell durations*, and thus of subjects in this sample. There is an order of these *single-spell durations* (i.e. $t_1 < t_2 < ... < t_N$). This order is present in the set $R(t_j)$. This is the set of *single-spell durations* that have transitioned at the order j, and that have not transitioned at the order j excluding those that already transitioned before the order j. Mathematically, this is $R(t_j) = \{l : t_l \geq t_j\}$.

The nonparametric Likelihood function for this study is

$$L(\zeta_1, \zeta_2, ..., \zeta_N; \hat{\beta}_y, \hat{\beta}_w) = \prod_{j=1}^{N} \left((\zeta_j^{-e^{\hat{\beta}_y^T y_j + \hat{\beta}_w^T w_j(t)}} - 1) \prod_{m \in R(t_j)} \zeta_j^{-e^{\hat{\beta}_y^T y_m + \hat{\beta}_w^T w_m(t)}} \right) \tag{6.11}$$

The $1 - \zeta_j$ are the baseline hazard rate contributions at the order j. This nonparametric Likelihood function allow the estimation of the *Survivor function*, and the contributions may be used to obtain a kernel smoothing function of the *hazard rate function.* This estimator is similar to the Kaplan-Meier estimator, but adjusted for the value of relative hazard's covariates.

The models are estimated using STATA (Cleves et al., 2010). The plots are also obtained using STATA (Mitchell, 2008).

6.4 Discussion and results

Table 6.1 presents the estimates of the relative hazard functions of the models. There are two types of models: Fixed Thresholds, and Moving Thresholds. The Fixed Thresholds models assume subjects have a rigid expectation with regards to their travel times, and also have a rigid magnitude of the travel time variability. The Moving Thresholds models assume subjects continuously update their margins based on previous past trips. Subjects classify their experienced trips whether they fall into the margins (i.e. *regular trip*), fall above the margins (i.e. *late trip*), and below the margins (*early trip*). Readers should refer to see section 6.3.3 for details. In addition, both Fixed Thresholds and Moving Thresholds assume that subjects have a moving set of travel times. The moving set of travel times refer to which travel times of past trips the subjects are able to recall. Furthermore, readers should remember that the dependent variable of the model is the *single-spell duration* (as defined in section 6.3) of the subjects in their current bridge until they decide to switch to another bridge or until their GPS devices are retrieved from their vehicles.

For the models (Fixed Thresholds and Moving Thresholds), several combinations of distinct levels of standard deviations (or moving standard deviations) from the mean (or moving mean), and of past trips in the moving set of travel times were tested. The results that were statistically significant are summarized in table 6.1. All the models find that the number of past trips for classifying *early trips* are less than 4 past trips. In contrast, the number of past trips for classifying *late trips* is greater than 3 past trips, and most of the time its value was found to be 6 past trips. This indicates that subjects were found to recall further in time travel experiences of greater travel times with respect to the mean (or moving mean) in comparison to travel experiences of smaller travel times with respect to the mean (or moving mean). In addition, 0.5 standard deviation (or moving standard deviation) from the mean (or moving mean) were found to be statistically significant at least 5% level for the margins classifying *early trips.* For the margins classifying *late trips*, 1 standard deviation (or moving standard deviations) from the mean (or moving mean) were found to be statistically significant at least 5%. This implies subjects have asymmetric margins for classifying a trip to be late or early. Also, subjects consider trips not too far below the mean (or moving mean) as early, but trips farther above the mean are considered as late. Thus, subjects tolerate travel times that are above, but close to the mean (or moving mean). Moreover, the signs indicate that the subjects are more likely to leave the *current bridge choice* if the number of *late trips* increases, and more likely to stay in the *current bridge choice* if the number of *early trips* increases. Readers should remember that Moving Thresholds models use moving means, and moving standard deviation, and the Fixed Threshold models use mean, and standard deviation. On the other hand, the addition of variables representing travel times of the subjects from past bridges were not found statistically significant. Thus, the question is whether subjects have a clean slate, or the inclusion of these variables requires a better hypothesis of how travelers see past time. It was previously discussed (see section 6.3.3 and section 6.2) that travelers may eventually forget about their past travel experience on other bridges. It is also plausible that travelers may change to other bridges, because they know of their existence, and not necessarily because they remember their previous travel experience. An interesting result of these variables, albeit not statistically significant, is that the sign is negative for the first week, and it is positive for the rest of the *single-spell duration*. This means that subjects with higher travel times on their past bridges are likely to stay in their *current bridge choice* during the first week. Similarly, traditional travel time measures (see section 2.4.1 and section 2.4.2) from the Centrality-Dispersion framework, and the Scheduling under uncertainty framework were not found statistically significant. Thus, we must question whether these frameworks are only reflecting the aggregate responses of subjects (i.e. static assumption), and are not able to capture the dynamics across the responses of subjects (see section 6.2).

Other important results are that the previous knowledge (i.e. before September 18th 2008) of other bridges is statistically significant at 1%. This previous knowledge (or past bridge diversity) has a positive sign indicating that subject familiar with other bridges are more susceptible to leave their *current bridge choice*. The ratio of bridge distances between the most used past bridge, and the *current bridge choice* indicates that subjects are influenced (statistically significant at 5%) by travel distance. Subjects prefer bridges that are closer for their home to work trips. Subjects are more susceptible to leave the *current bridge choice* is they know of a previous past bridge that is closer to them. In addition, subjects that indicated in the surveys fear of bridges are more susceptible (statistically significant at least 10%) to leave *current bridge choice* as long as the *current bridge choice* is the I-35W bridge or other bridges in the vicinity (i.e. Washington Ave bridge, and 10th St bridge). Lastly, the income levels of the subjects were found statistically significant at least 5% (see section 6.3.3). The sign of these variables is negative, and thus indicating a reluctance to leave their *current bridge choice*. These variables require further research to identify the reason behind them. It is hypothesized that perhaps the type of jobs may play a role.

Finally, the goodness of fit analysis indicate that the Cox PH models' proportional hazard hypothesis cannot be rejected (i.e. Schoenfeld p-values), and the Deviance plots indicate for most models that there is no pattern, except a slight pattern for the Moving Thresholds models. Moreover, the Cox PH model preferred according to Akaike Information Criteria, and Bayesian Information criteria is the Fixed Thresholds 2 model.

Figures 6.3a, 6.3b, and 6.3c present the estimates of the baseline cumulative hazard, baseline hazard rate function, and baseline survivor function for the Fixed Thresholds 2 model. The baseline cumulative hazard indicates that there is rapid growth in the susceptibility to leave the *current bridge choice*, and that eventually it flattens. This rapid growth happens before 40 days after the date September 18th 2008. This agrees with the baseline hazard rate function that most of the susceptibility to leave the *current bridge choice* occurs before 40 days after the date September 18th 2008. In addition, the susceptibility of leaving the *current bridge choice* increases until 20 days after September 18th 2008, and eventually the susceptibility decreases. This agrees with the baseline survivor function that indicates a sharp drop in the survival probability by 20 days after September 18th 2008, and a smoother drop between 20 days after September 18th 2008, and 40 days after September 18th 2008.

6.5 Conclusion

Currently, models that incorporate travel time reliability are *static models* and *Random Utility models* (Ben-Akiva and Lerman, 1985, Train, 2009, Ortuzar and Willumsen, 2011). The models are *static* because they only consider trade-offs of statistical measures on travel time distributions of a set of alternatives (e.g. routes, bridges). The travel time distributions may be hypothetical (stated preference data) or day-to-day travel times (revealed preference data). In stated preference studies, the travel time distribution of each alternative is generated by the researchers. In addition, the researchers are familiar with the value of different statistical measures (e.g. mean, median), based on the previous theoretical frameworks (see section 2.2), of the travel time distribution of each alternative. Thus, subjects simply observe a set of attributes of each alternative that are abstractions of the concept of travel time reliability/variability, and of expected travel time. These abstractions depend on different presentations (e.g. numbers and/or visual aid), and in cases may include the travel time distributions of the alternatives such as in the case of histograms in the stated choice questions. On the other hand, revealed preference studies produce travel time distributions for an alternative by aggregating travel times of different days of the alternative. Statistical measures are computed on these day-to-day travel times, and it is assumed that these measures represent the experience of the subjects with regards to the travel time reliability/variability, and the expected travel time for each alternative. In essence, the dynamic behavior of the travelers have been neglected to favor an assumption that subjects, at the aggregate level, settle for a particular alternative. However, this assumption neglects that subjects may be exercising their decisions at a day to day level. In addition, the subjects may only consider a subset of the day to day travel times for each alternative, and that this subset of day to day travel times is updated every few days. In other words, subjects may consider adding to the set day-to-day travel times according to their perception, and the subjects may discard day-to-day travel times according to their limited memory. Furthermore, the models are *Random Utility models*, and thus they are bound by the assumption that travelers as utility-maximizers know the expected travel time, and the travel time variability of each alternative. Travelers are *informed* with regards to the attributes related to travel time (i.e. means, standard deviation) for each alternative.

This is a fundamental assumption in *Random Utility models* based on the theory of rational choice presented in Domencich and McFadden (1975). It is possible to include only the travel times known for certain alternatives, but this implicitly assumes that the travel times of the alternatives, that excludes them, are of the highly unlikely value of zero. This is also a questionable assumption. In fact, questions arise about whether a subject truly knows of the travel time of his other alternatives, or a subject only knows the travel time obtained from past accumulated experience. In addition, questions arise about the modus operandi of a subject to choose whether to stay or abandon the current chosen alternative based on its day to day travel times. In light of this discussion, researchers must ask whether stated preference studies are presenting a decision-making situation that is *realistic*, and whether revealed preference studies with travel time distributions obtained by aggregating day to day travel times are *simplistic*.

In this study, the authors have tried to uncover the dynamic behavior of subjects by using GPS data, and using an alternative modeling approach (*duration models*) to the *Random Utility models*. Cox PH models are fitted to the *single-spell durations* of travelers after a new bridge replacing the collapsed I-35W bridge in Minneapolis opened to the public. Several key difficulties are identified: travelers *perceive* the travel times at the end of each of their trips and travelers *recall* travel times from past trips. Thus, two important concerns are: the selection of day to day travel time by travelers due to their perception; and the ability of the travelers to recall the previously selected day to day travel times. Furthermore, two types of models are proposed: Fixed Thresholds models, and Moving Thresholds models. The Fixed Thresholds models assume subjects have a rigid expectation with regards to their travel times, and also have a rigid magnitude of the travel time variability. The Moving Thresholds models assume subjects continuously update their margins based on previous past trips. Subjects classify their experienced trips whether they fall into the margins (i.e. *regular trip*), fall above the margins (i.e. *late trip*), and below the margins (*early trip*).

The primary results indicate that both the Fixed Thresholds, and the Moving Thresholds models are found to capture the dynamics of the data, but the the Fixed Thresholds should be preferred. In addition, the *late trips* are more persistent in the subjects' travel time moving sets in comparison to the subjects' *early trips.* It is also found that both margins to classify whether a trip is an *early trip* or a *late trip* are asymmetric. Subjects are tolerant to small increases above the mean in their travel experiences. The secondary results indicate that subjects' perception (i.e. fear, travel distance) of the alternatives also influence their decision to abandon the chosen route.

Future research should look into the connection between Fixed Thresholds and Moving Thresholds models, because it is likely that subjects are indeed continuously updating their margins, but this updating process is highly dependent on the subjects' perception of travel time. In addition, the incorporation of travel time measures from past travel experiences in other alternatives still requires more research. It is important to ascertain when subjects eventually forget (if they forget at all) about past travel experience, and also whether is more important that subjects are familiar with other alternatives or must subjects also have travel experience with other alternatives. Lastly, there is significant interest to revisit the data with a more advanced duration model that is able to capture these Thresholds-based models in a recurrent spell duration setting.

Table 6.1: Cox Proportional Hazard Models

Variables	Fixed Thresholds 1[a]	Fixed Thresholds 2[a]	Fixed Thresholds 3[a]	Fixed Thresholds 4[a]	Moving Thresholds 1[b]	Moving Thresholds 2[b]
Relative Hazard: $\phi(\mathbf{y}, \mathbf{w(t)}; \beta_\mathbf{y}, \beta_\mathbf{w}) = \mathbf{e}^{\beta_\mathbf{y}^\mathbf{T}\mathbf{y}+\beta_\mathbf{w}^\mathbf{T}\mathbf{w(t)}}$ **(see section 6.3)**	**Early Trips:** Mean - 0.5SD; 3 past trips **Late Trips:** Mean + 1SD; 3 past trips	**Early Trips:** Mean - 0.5SD; 4 past trips **Late Trips:** Mean + 1SD; 8 past trips	**Early Trips:** Mean - 0.5SD; 2 past trips **Late Trips:** Mean + 1SD; 6 past trips	**Early Trips:** Mean - 0.5SD; 3 past trips **Late Trips:** Mean + 1SD; 6 past trips	**Early Trips:** M. Mean - 0.5 M. SD; 3 past trips **Late Trips:** M. Mean + 1M. SD; 4 past trips	**Early Trips:** M. Mean - 0.5 M. SD; 3 past trips **Late Trips:** M. Mean + 1M. SD; 6 past trips
	Estimates (T-Stats)	**Estimates (T-Stats)**	**Estimates (T-Stats)**	**Estimates (T-Stats)**	**Estimates (T-Stats)**	**Estimates (T-Stats)**
Past bridge diversity - y_1	1.61 (5.05) ***	1.64 (5.11) ***	1.62 (4.96) ***	1.61 (5.01) ***	1.45 (4.69) ***	1.63 (5.07) ***
Gender (Male = 1; 0 = Female) - y_2	0.391 (0.79)	0.447 (0.89)	0.416 (0.85)	0.472 (9.96)	0.377 (0.78)	0.385 (0.77)
Income $(\$50,000, \$99,999]$ **(1 = In; 0 = Out) -** y_3	-1.18 (-2.43) **	-1.11 (-2.32) **	-1.21 (-2.47) **	-1.18 (-2.43) **	-0.80 (-1.81) *	-0.80 (-1.80) *
Income $(\$100,000, \infty+)]$ **(1 = In; 0 = Out) -** y_3	-1.74 (-2.77) ***	-1.86 (-2.88) ***	-1.73 (-2.82) ***	-1.75 (-2.84) ***	-1.46 (-2.45) **	-1.33 (-2.17) **
Ratio of bridge distances - y_4 **(Current bridge choice to previous bridge choice)**	1.71 (1.99) **	1.80 (2.09) **	1.45 (1.69) *	1.66 (1.93) **	1.69 (2.00) **	1.64 (1.84) **
Fear of driving on the I-35W bridge - y_5 **and other bridges in the vicinity**	0.72 (1.87) **	0.80 (2.03) **	0.76 (1.92) **	0.79 (2.01) **	0.69 (1.78) *	0.65 (1.67) *
Fixed Early trips - $w_1(t)$ **Mean** $-\gamma$ **Std. Dev** δ **past trips**	-1.75 (-2.84) *** $M - 0.5SD$ 3 past trips	-1.93 (-2.79) *** $M - 0.5SD$ 4 past trips	-1.62 (-2.87) *** $M - 0.5SD$ 2 past trips	-1.81 (-2.89) *** $M - 0.5SD$ 3 past trips		
Moving Early trips - $w_2(t)$ **M. Mean** $-\gamma$ **M. Std. Dev** δ **past trips**					-2.55 (-2.43) ** $MM - 0.5MSD$ 3 past trips	-1.83 (-1.97) ** $MM - 0.5MSD$ 3 past trips
Fixed Late trips - $w_1(t)$ **Mean** $+\gamma$ **Std. Dev** δ **past trips**	2.06 (1.69) * $M + 1SD$ 3 past trips	5.82 (2.61) *** $M + 1SD$ 8 past trips	4.48 (2.42) ** $M + 1SD$ 6 past trips	4.29 (2.40) ** $M + 1SD$ 6 past trips		
Moving Late trips - $w_2(t)$ **M. Mean** $+\gamma$ **M. Std. Dev** δ **past trips**					4.04 (1.78) * $MM + 1MSD$ 4 past trips	2.43 (2.26) ** $MM + 0.5MSD$ 6 past trips
Median of previous bridge choice - $w_3(t)$	0.044 (1.23)	0.032 (0.96)	0.046 (1.33)	0.042 (1.21)	0.0090 (0.32)	0.0061 (0.21)
Median of previous bridge choice - $w_3(t)$ **(5 days or less)**	-0.13 (-1.40)	-0.12 (-1.25)	-0.11 (-1.13)	-0.13 (-1.36)	-0.10 (-1.09)	-0.11 (-1.12)
Partial Log-Likelihood $ll_{\hat{\beta}}$	-106.9997	-105.5034	-105.9789	-105.8765	-109.1614	-108.3049
Akaike Information Criterion AIC	233.9994	231.0068	231.9577	231.7531	238.3228	236.6099
Bayesian Information Criterion BIC	281.2737	278.2812	279.232	279.0274	285.5971	283.8842
Schoenfeld p-value SCH	0.3198	0.2963	0.4248	0.3958	0.2198	0.2055
Deviance plot D_p	6.2a	6.2b	6.2c	6.2d	6.2e	6.2f
Number of observations	835	835	835	835	835	835
Number of subjects	65	65	65	65	65	65

* is 10% significance level, ** is 5% significance level, *** is 1% significance level

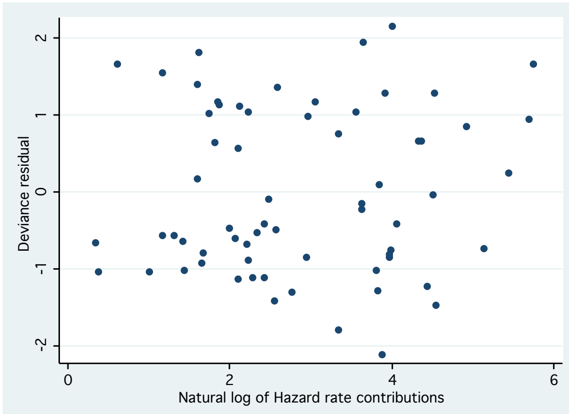
[a] There two rigid margins that classify trips (in a moving set) as: early trips, and late trips. The moving set only contain previous trips to the current trip. (see section 6.3.3).

[b] There two moving margins that classify trips as: early trips, and late trips. The moving set only contain previous trips to the current trip. (see section 6.3.3).

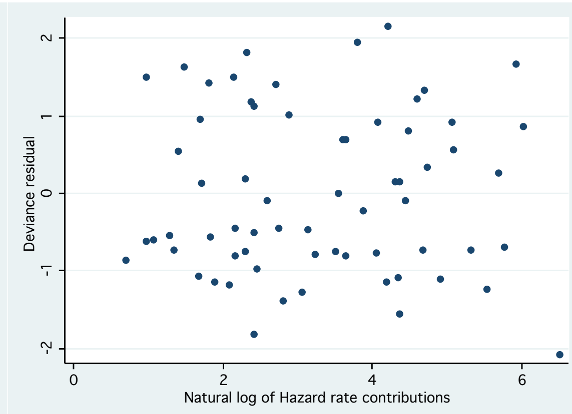
See the section 6.3 for details on the econometric models.

Figure 6.2: Deviance residuals plots

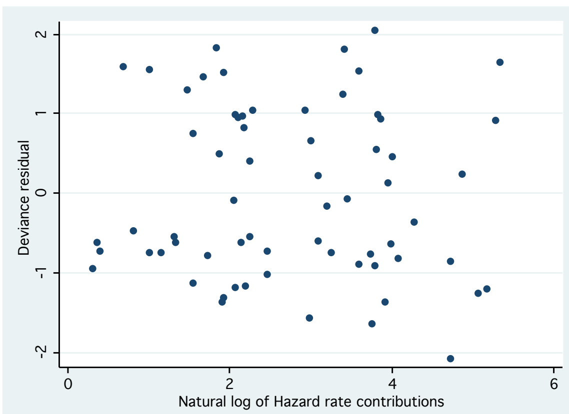
(a) Deviance Plot - Fixed Thresholds 1



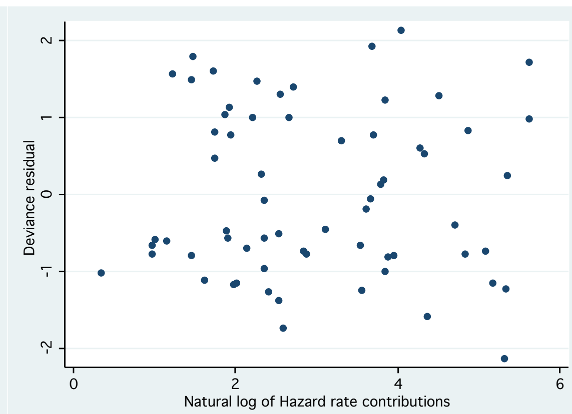
(b) Deviance Plot - Fixed Thresholds 2



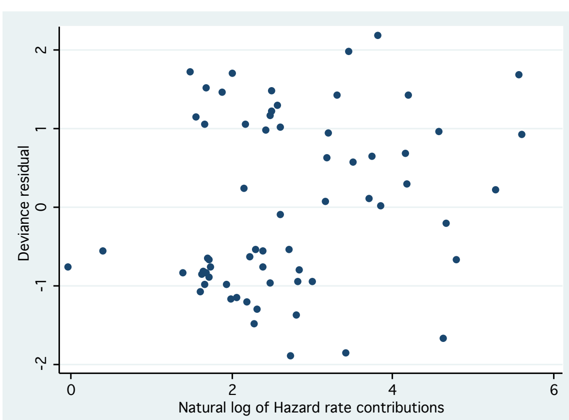
(c) Deviance Plot - Fixed Thresholds 3



(d) Deviance Plot - Fixed Thresholds 4



(e) Deviance Plot - Moving Thresholds 1



(f) Deviance Plot - Moving Thresholds 2

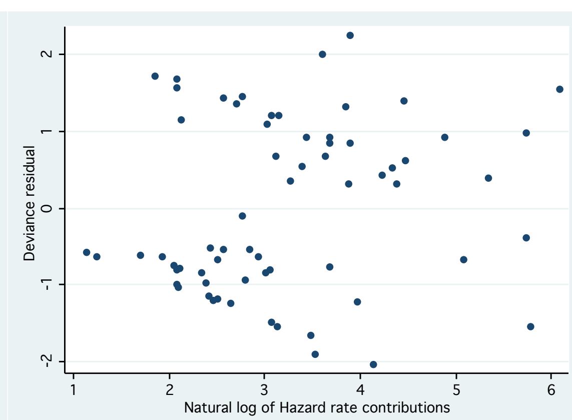
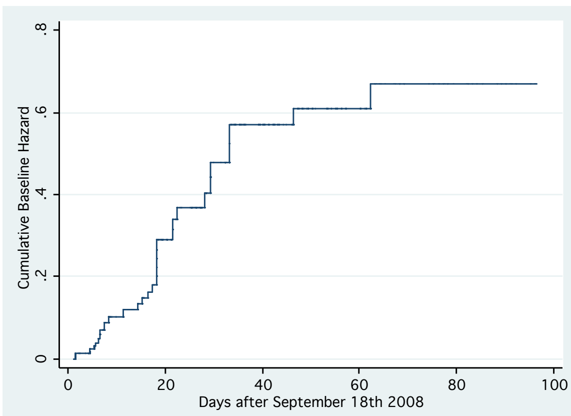
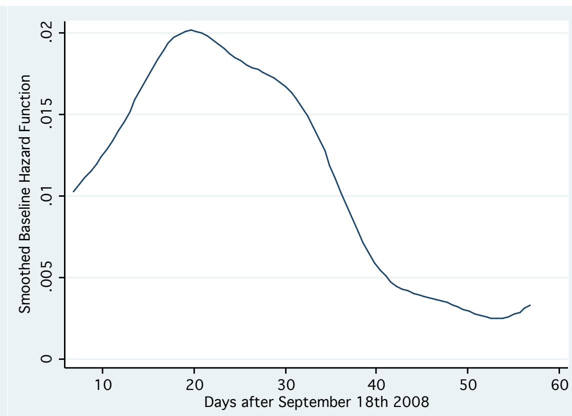


Figure 6.3: Baseline functions

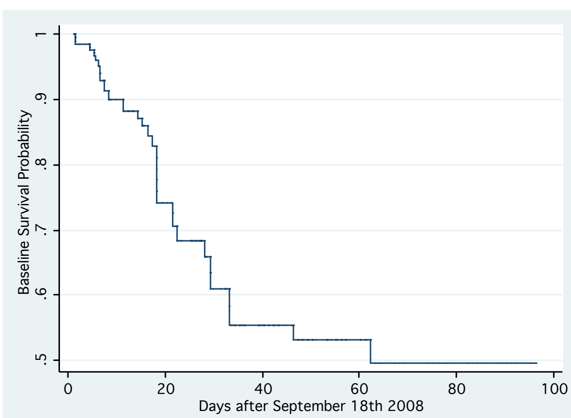
(a) Baseline cumulative hazard - Fixed Thresholds 2



(b) Smoothed baseline hazard function - Fixed Thresholds 2



(c) Baseline survivor function - Fixed Thresholds 2



Chapter 7

Conclusions

There has been little effort in studying empirically the causes and consequences of *perception errors* in travelers despite that there are studies (Levinson et al., 2004, 2006, Peer et al., 2010) indicating that travelers overestimate or underestimate the actual travel times they experience.

In this thesis, the factors causing commuters' perception error are investigated in chapter 4. hese factors are included in econometric models to study their influence on travel time perception, and also ascertain which of these factors lead to overestimation or underestimation of travel times. These econometric models, described in section 4.2, are fitted using three functional forms (linear, quadratic, and Cobb-Douglas) on data collected of commuters recruited from a previous research study in the Minneapolis-St. Paul region (Zhu, 2010, Carrion and Levinson, 2012a). This data (surveys, and Global Positioning System [GPS] points) consists of work trips (from home to work, and from work to home) of subjects. For these work trips, the subjects' self-reported travel times, and the subjects' travel times measured by GPS devices were collected. The results continue to highlight the need to further study the perception of travel time, and to acknowledge its influence in travelers' decisions. Furthermore, the factors are related into four categories based on time perception research in psychology: temporal relevance; temporal uncertainty, and temporal expectancies; task complexity, and absorption, and attentional deployment; and affective elements. This allows the factors to be compared to similar research of time perception in psychology.

The consequences of commuters' perception error has been largely ignored. This same data, described previously, is used to investigate consequences of the presence of *perception error* in the valuation of travel time for economic analyses, and in the dynamic behavior of commuters for route choice modeling.

The influence of commuters' perception error in the valuation of travel time is investigated by estimating random utility models on the previously described data. The results indicate that the subjects value travel time reliability more than travel time savings (i.e. reliability ratios greater than 1) in the econometric models with self-reported travel times. In contrast, subjects value travel time savings more than travel time reliability (i.e. reliability ratios smaller than 1) in the econometric models with travel times as measured by GPS devices. Furthermore, the models with self reported travel times are found to statistically fit the data better in comparison to the models with measured travel times. Thus, the reliability ratio (a important quantity in economic analyses) is inflated or deflated according to perception error.

The behavior of commuters with regards to day to day route choices is investigated using Cox Proportional Hazard models on the previously described data. Several key difficulties are identified: travelers *perceive* the travel times at the end of each of their trips; and travelers *recall* travel times from past trips. Two types of models are proposed: Fixed Thresholds models, and Moving Thresholds models. The Fixed Thresholds models assume subjects have a rigid expectation with regards to their travel times, and also have a rigid magnitude of the travel time variability. The Moving Thresholds models assume subjects continuously update their margins based on previous past trips. Also, moving sets of travel times are specified to identify which travel times of past trips the subjects are able to recall. The primary results indicate that the subjects react to day-to-day travel times on a specific route according to thresholds. These thresholds help discriminate whether a travel time is within an acceptable margin or not, and travelers may decide to abandon the chosen route depending on the frequency of travel times within acceptable margins. The secondary results indicate that subjects previous experience, and perception of the alternatives also influence their decision to abandon the chosen route.

In summary, factors causing the perception error of commuters are identified, and are related into four categories based on time perception research in psychology. This allows the factors to be compared to similar research of time perception in psychology. Furthermore, the consequences of perception errors are investigated in the valuation of travel time, and in route choice modeling. The results indicate that perception error indeed influences the valuation of travel time. In addition, the dynamics of travelers' route choices may be influenced by perception error, and memory recollection. Thus, the modeling of travel decisions must account for perception error. This is an important research topic as many analyses (e.g. economic, planning), and models (e.g. economic, traffic) in transportation continue to ignore that travelers are executing decisions according to their own divergent views of the actual travel time distributions.

Bibliography

2006-2008 American Community Survey 3-Year Estimates, Minneapolis-St. Paul-Bloomington, MN-WI Metropolitan Statistical Area, Retrieved November 25, 2009.

Abdel-Aty, M., R. Kitamura, and P. Jovanis (1997). Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. *Transportation Research Part C 5*(1), 39–50.

Abrantes, P. and M. Wardman (2011). Meta-analysis of UK values of travel time: An update. *Transportation Research Part A 45*, 1–17.

Allan, L. (1979). The perception of time. *Perception & Psychophysics 26*(5), 340–354.

Angrilli, A., P. Cherubini, A. Pavese, and S. Manfredini (1997). The influence of affective factors on time perception. *Perception & Psychophysics 59*, 972–982.

Arnott, R., A. De Palma, and R. Lindsey (1990). Economics of a Bottleneck. *Journal of Urban Economics 27*(1), 111–130.

Arnott, R., A. De Palma, and R. Lindsey (1993). A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand. *The American Economic Review 83*(1), 161–179.

Arnott, R., A. De Palma, and R. Lindsey (1994). The welfare effects of congestion tolls with heterogeneous commuters. *Journal of Transport Economics and Policy 28*, 139–161.

Asensio, J. and A. Matas (2008). Commuters' valuation of travel time variability. *Transportation Research Part E: Logistics and Transportation Review 44*(6), 1074–1085.

ATOC. *Passenger Demand Forecasting Handbook*. Association of Train Operating Companies, London (UK).

Bartram, D. J. (1980). Comprehending spatial information: the relative efficiency of different methods of presenting information about bus routes. *Journal of Applied Psychology 65*, 103–110.

Bates, J., J. Polak, P. Jones, and A. Cook (2001). The valuation of reliability for personal travel. *Transportation Research Part E 37*, 191–229.

Batley, R. and J. N. Ibanez (2009). Randomness in preferences, outcomes and tastes; an application to journey time risk. In *Proceedings of the International Choice Modelling Conference, Harrogate, UK.*

Becker, G. (1965). A Theory of the Allocation of Time. *The Economic Journal 75*(229), 493–517.

Ben-Akiva, M., M. Bergman, A. Daly, and R. Ramaswamy (1984). Modelling inter urban route choice behaviour. In *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory, Delft, the Netherland*, pp. 299–330.

Ben-Akiva, M. and S. Lerman (1985). *Discrete choice analysis: theory and application to travel demand*. MIT Press.

Bhat, C. and R. Sardesai (2006). The impact of stop-making and travel time reliability on commute mode choice. *Transportation Research Part B: Methodological 40*(9), 709–730.

Blaauw, G. J. (1982). Driving experience and tasks demands in simulator and instrumented car: a validation study. *Human Factors 24*, 473–486.

Black, I. and J. Towriss (1993). Demand effects of travel time reliability. Technical report, Centre for Logistics and Transportation, Craneld Institute of Technology.

Block, R. and D. Zakay (1996). Models of psychological time revisited. In H. Helfrich (Ed.), *Time and mind*, pp. 171–195. Hogrefe and Huber.

Boltz, M. (1993). Time estimation and expectancies. *Memory and Cognition 21*, 853–863.

Bovy, P. and E. Stern (1990). *Route Choice: Wayfinding in Transport Networks*. Netherlands: Kluwer Academic Publishers.

Bruzelius, N. (1979). *The Value of Travel Time: theory and measurement*. Croom Helm London.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge Univ. Press.

Cameron, A. C. and P. K. Trivedi (2010). *Microeconometrics using STATA* (Revised ed.). Stata Press.

Carrion, C. and D. Levinson (2012a). A Model of Bridge Choice Across the Mississippi River in Minneapolis. In D. Levinson, H. Liu, and M. Bell (Eds.), *Network Reliability in Practice: Selected papers from the fourth international symposium on transportation network reliability*, Chapter 8, pp. 115–129. Springer.

Carrion, C. and D. Levinson (2012b). Valuation of travel time reliability from a GPS-based experimental design. *Transportation Research Part C (In Press)*.

Carrion, C. and D. Levinson (2012c). Value of Reliability: A review of the current evidence. *Transportation Research Part A 46*(4), 720–741.

Chang, M. F. and R. Herman (1978). An attempt to characterize traffic in metropolitan areas. *Transportation Science 12*, 59–79.

Chang, X. and P. Stopher (1981). Defining the perceived attributes of travel modes for urban work trips. *Transportation Planning and Technology 7*, 55–65.

Cleves, M., R. Gutierrez, W. Gould, and Y. Marchenko (2010). *An Introduction to Survival Analysis using STATA* (3rd ed.). Stata Press.

Collett, D. (2003). *Modelling Survival Data in Medical Research* (2nd ed.). Chapman and Hall.

Cook, A., P. Jones, J. Bates, J. Polak, and M. Haigh (1999). Improved methods of representing travel time reliability in SP experiments. In *Proceedings 27th European Transport Forum, PTRC, London.*

Copley, G., P. Murphy, and D. Pearce (2002). Understanding and valuing journey time variability. In *Proceedings of the European Transport Conference, Cambridge. PTRC, London.*

Cramer, J. (1986). *Econometric applications of Maximum Likelihood methods.* Cambridge Univ. Press.

Daganzo, C. (1985). The Uniqueness of a Time-dependent Equilibrium Distribution of Arrivals at a Single Bottleneck. *Transportation Science 19*(1), 29–37.

Daganzo, C. (1995). A Pareto Optimum Congestion Reduction Scheme. *Transportation Research Part B: Methodological 29B*(2), 139–154.

Daganzo, C. and R. Garcia (2000). A Pareto improving strategy for the time-dependent morning commute problem. *Transportation Science 34*, 303–311.

Daganzo, C. F. (2007). *Fundamentals of Transportation and Traffic Operations.* Emerald Group.

DeSerpa, A. (1971). A theory of the economics of time. *The Economic Journal 81*(324), 828–846.

D'Este, G. (1986). Route preference surveys and possible network flows. In *8th Conference of Australian Institute of Transportation Research, Australia.*

Domencich, T. and D. McFadden (1975). *Urban Travel Demand-A Behavioral Analysis.* North Holland.

Duffell, J. R. and A. Kalombaris (1988). Empirical studies of car driver route choice in Hertfordshire. *Traffic Engineering and Control*, 398–408.

Fosgerau, M. and L. Engelson (2011). The value of travel time variance. *Transportation Research Part B 45*(1), 1–8.

Fosgerau, M. and A. Karlstrom (2010). The value of reliability. *Transportation Research Part B 44*(1), 38–49.

Fox, R., P. Bradbury, and I. Hampton (1967). Time judgment and body temperature. *Journal of Experimental Psychology 75*, 88–96.

Garcia, R. (1999). *A Pareto improving strategy for the time-dependent morning commute problem.* Ph. D. thesis, University of California, Berkeley (USA).

Gaver, D. (1968). Headstart strategies for combating congestion. *Transportation Science 2*(2), 172–181.

Ghosh, A. (2001). *Valuing time and reliability: commuters' mode choice from a real time congestion pricing experiment.* Ph. D. thesis, University of California, Irvine (USA).

Glicksohn, J., a. M. B. and E. Pavell (1991-1992). Imagination, absorption and subjective time estimation. *Imagination, Cognition and Personality 11*, 167–176.

Godley, S. T., T. Triggs, and B. Fildes (2002). Driving Simulator validation for speed research. *Accident Analysis and Prevention 34*, 589–600.

Goldin, S. E. and P. W. Thorndyke (1982). Simulating navigation for spatial knowledge acquisition. *Human Factors 24*, 457–471.

Golledge, R. (1992). Place Recognition and Wayfinding: Making Sense of Space. *Geoforum 23*, 199–214.

Golledge, R. (1999). Human wayfinding and cognitive maps. In *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*, pp. 5–45. John Hopkins.

Golledge, R. and R. Stimson (1997). *Spatial behavior.* Guilford Press, New York.

Grambsch, P. and T. Therneau (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika 81*, 515–526.

Greene, W. (2012). *Econometric Analysis* (7th ed.). Prentice-Hall.

Haefner, L. E. and L. V. Dickinson (1974). Preliminary Analysis of Disaggregate Modeling in Route Choice. *Transportation Research Record* (527), 66–72.

Hamerslag, R. (1981). Investigation into factors affecting the route choice in Rijnstreek-West with the aid of a disaggregate logit model. *Transportation 10*(4), 373–391.

Hazell, P. and R. Norton (1986). *Mathematical Programming for Economic Analysis in Aggriculture.* Macmillan, London.

Hensher, D. (1994). Stated preference analysis of travel choices: the state of practice. *Transportation 21*(2), 107–133.

Hensher, D. (2001). The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications. *Transportation 28*, 101–118.

Hensher, D. (2010). Hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B 44*, 735–752.

Hollander, Y. (2006). Direct versus indirect models for the effects of unreliability. *Transportation Research Part A: Policy and Practice 40*(9), 699–711.

Jackson, W. and J. Jucker (1982). An empirical study of travel time variability and travel choice behavior. *Transportation Science 16*(4), 460–475.

Jara-Diaz, S. (2000). Allocation and valuation of travel time savings. In D. Hensher and K. Button (Eds.), *Handbook of transport modelling*, pp. 304–319. Pergamon Press.

Jara-Diaz, S. (2007). *Transport Economic Theory*. Emerald Group.

Jenelius, E. (2012). The value of travel time variability with trip chains, flexible scheduling and correlated travel times. *Transportation Research Part B 46*(6), 762–780.

Jenelius, E., L.-G. Mattsson, and D. Levinson (2011). Traveler delay costs and value of time with trip chains, flexible activity scheduling and information. *Transportation Research Part B 45*, 789–807.

Johnston, J. and J. DiNardo (1997). *Econometric methods*. McGraw-Hill.

Jones, M. and M. Boltz (1989). Dynamic attending and responses to time. *Psychological Review 96*, 459–491.

Kahneman, D. and A. Tversky (1973). On the Psychology of Prediction. *Psychological Review 80*(4), 237–251.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47*(2), 263–292.

Kalbfleisch, J. and R. Prentice (2002). *The statistical analysis of failure time data* (2nd ed.). Wiley.

Kansky, K. J. (1967). Travel Patterns of Urban Residents. *Transportation Science 1*(4), 261–285.

Khattak, A., F. Koppelman, and J. Schofer (1993). Stated Preference for investigating commuters' diversion propensity. *Transportation 20*, 107–127.

Knight, T. (1974). An approach to the evaluation of changes in travel unreliability: A safety margin hypothesis. *Transportation 3*, 393–408.

Koskenoja, P. (1996). *The Effect of Unreliable Commuting Time on Commuter Preferences*. Ph. D. thesis, University of California, Irvine (USA).

Laih, C. (1994). Queueing at a bottleneck with single-and multi-step tolls. *Transportation Research Part A: Policy and Practice 28A*, 197–208.

Lam, T. and K. Small (2001). The value of time and reliability: measurements from a value pricing experiment. *Transportation Research Part E 37*, 235–251.

Langer, J., S. Wapner, and H. Werner (1961). The effect of danger upon the experience of time. *American Journal of Psychology 74*, 94–97.

Leiser, D. and E. Stern (1988). Determinants of subjective time estimates in simulated urban driving. *Transportation Research A 22*, 175–182.

Levinson, D., K. Harder, J. Bloomfield, and K. Carlson (2006). Waiting Tolerance: Ramp Delay vs. Freeway Congestion. *Transportation Research part F: Traffic Psychology and Behaviour 9*(1), 1–13.

Levinson, D., K. Harder, J. Bloomfield, and K. Winiarczyk (2004). Weighting Waiting: Evaluating the Perception of In-Vehicle Travel Time Under Moving and Stopped Conditions. *Transportation Research Record: Journal of the Transportation Research Board 1898*, 61–68.

Li, H. (2004). *Investigating Morning Commute Route Choice Behavior Using Global Positioning Systems and Multi-day Travel Data*. Ph. D. thesis, Georgia Institute of Technology (USA).

Li, H., R. Guensler, and J. Ogle (2005). Analysis of Morning Commute Route Choice Patterns Using Global Positioning System-Based Vehicle Activity Data. *Transportation Research Record: Journal of the Transportation Research Board 1926*, 162–170.

Li, H., R. Guensler, J. Ogle, and J. Wang (2004). Using global positioning system data to understand day-to-day dynamics of morning commute behavior. *Transportation Research Record: Journal of the Transportation Research Board 1895*, 78–84.

Li, Z., D. Hensher, and J. Rose (2010). Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transportation Research Part E 46*, 384–403.

Li, Z., A. Tirachini, and D. Hensher (2012). Embedding Risk Attitudes in a Scheduling Model: Application to the Study of Commuting Departure Time. *Transportation Science 46*(2), 170–188.

Liu, H., X. He, and W. Recker (2007). Estimation of the time-dependency of values of travel time and its reliability from loop detector data. *Transportation Research Part B: Methodological 41*(4), 448–461.

Liu, H., W. Recker, and A. Chen (2004). Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data. *Transportation Research Part A 38*, 435–453.

Louviere, J., D. Hensher, and J. Swait (2000). *Stated choice methods: analysis and applications*. Cambridge University Press.

Madalina, S. (2011). *Cognitive Mechanisms involved in the subjective time perception*. Ph. D. thesis, Babes-Bolyai University (Romania).

Mahmassani, H. S. and R. Herman (1989). Interactive experiments for the study of tripmaker behaviour dynamics un congested commuting systems. In *Oxford Conference on Travel Behaviour*.

Markowitz, H. (1999). The early history of portfolio theory: 1600-1960. *Financial Analysis Journal 55*, 5–16.

Mas-Colell, A., M. Whinston, and J. Green (1995). *Microeconomic Theory.* Oxford University Press.

Michaels, R. D. (1966). Attitudes of Drivers Toward Alternative Highways and Their Relation to Route Choice. *Highway Research Record* (122), 50–74.

Michea, A. and J. Polak (2006). Modelling risky choice behaviour: Evaluating alternatives to expected utility theory. In *11th International Conference on Travel Behaviour Research, Kyoto.*

Mitchell, M. (2008). *A Visual Guide to STATA Graphics* (2nd ed.). Stata Press.

Newell, G. (1987). The morning commute for nonidentical travelers. *Transportation Science 21*(2), 74–88.

Nicholson, A. and Z. Du (1997). Degradable transportation systems: an integrated equilibrium model. *Transportation Research Part B 31*(3), 209–223.

Noland, R. and K. Small (1995). Travel-time uncertainty, departure time choice, and the cost of morning commutes. *Transportation Research Record 1493*, 150–158.

Ortuzar, J. and L. Willumsen (2011). *Modelling Transport* (4th ed.). Wiley.

Pal, A. (2004). Modeling of Commuter's Route Choice Behavior. Master's thesis, The University of Toledo (USA).

Parthasarathi, P. (2011). *Network Structure and Travel.* Ph. D. thesis, University of Minnesota, Twin Cities (USA).

Parthasarathi, P., D. Levinson, and H. Hochmair (2012). Network Structure and Travel Time Perception. In *Presented at the International Association of Travel Behavior Research Conference in Toronto.*

Pawitan, Y. (2001). *In All Likelihood: Statistical Modelling and Inference Using Likelihood.* Oxford University Press.

Peer, S. (2013). *The Economics of Trip scheduling, Travel time variability, and Traffic information.* Ph. D. thesis, VU University Amsterdam (Netherlands).

Peer, S., P. Koster, and E. Verhoef (2010). The perception of travel time variability. In *Conference Proceedings of 4th International Symposium on Tranportation Network Reliability in Minneapolis, MN (USA).*

Pells, S. (1987). *The evaluation of reductions in travel time variability.* Ph. D. thesis, University of Leeds (UK).

Polak, J. (1987). A more general model of individual departure time choice. In *PTRC Summer Annual Meeting, Proceedings of Seminar C.*

Polak, J. (1996). An overview of the recent literature on modelling the effects of travel time variability. Working Paper (London: Centre for Transport Studies, Imperial College).

Prashker, J. (1979). Direct Analysis of the perceived importance of attributes of reliability of travel modes in urban travel. *Transportation 8*, 329–346.

Schachter, S. and J. Singer (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review 69*, 379–399.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika 69*, 239–241.

Schwanen, T. and D. Ettema (2009). Coping with unreliable transportation when collecting children: Examining parents behavior with cumulative prospect theory. *Transportation Research Part A 43*, 511–525.

Scott, D. (1985). Driving simulator to aid road safety. *Automotive engineering 93*, 79–82.

Senbil, M. and R. Kitamura (2004). Reference points in commuter departure time choice: A prospect theoretic test of alternative decision frames. *Journal of Intelligent Transportation Systems 8*(1), 19–31.

Senna, L. (1994). The Influence of Travel Time Variability on the Value of Time. *Transportation 21*, 203–228.

Shires, J. and G. de Jong (2009). An international meta-analysis of value of travel time savings. *Evaluation and Program Planning 32*, 315–325.

Simon, H. (1997). *Models of bounded rationality: empirically grounded economic reason*, Volume 3. MIT Press.

Small, K. (1982). The Scheduling of Consumer Activities: Work Trips. *American Economic Review 72*(3), 467–479.

Small, K., R. Noland, X. Chu, and D. Lewis (1999). Valuation of Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation. Technical report, National Cooperative Highway Research Program Report 431.

Small, K., R. Noland, and P. Koskenoja (1995). Socio-economic Attributes and impacts of travel reliability: a stated preference approach. Technical report, California PATH Research Report, UCB-ITS-PRR-95-36.

Small, K. and E. Verhoef (2007). *The Economics of Urban Transportation.* Routledge, part of the Taylor & Francis Group.

Small, K., C. Winston, and J. Yan (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica 73*(4), 1367–1382.

Small, K., C. Winston, and J. Yan (2006). Differentiated road pricing, express lanes, and carpools: Exploiting heterogeneous preferences in policy design. *Brookings-Wharton Papers on Urban Affairs 7*, 53–96.

Takayama, A. (1993). *Analytical Methods in Economics.* University of Michigan Press.

TBI (2003). *2000 Travel Behavior Inventory home interview survey : data and methodology.* Metropolitan Council of the Twin Cities Area, St. Paul, Minnesota, USA.

Tellegen, A. and G. Atkinson (1974). Openness to absorbing and self-altering experiences (absorption), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology 83*, 268–277.

Thayer, S. and W. Schiff (1975). Eye-contact, facial expression, and the experience of time. *The Journal of Social Psychology 95*, 117–124.

Thomas, E. and W. Weaver (1975). Cognitive processing and time perception. *Perception & Psychophysics 17*, 363–367.

Tilahun, N. and D. Levinson (2009). *Travel Demand Management and Road User Pricing: Success, Failure and Feasibility.*, Chapter Unexpected delay and the cost of lateness on I-394 high occupancy/toll lanes., pp. 173–184. Ashgate Publishers.

Tilahun, N. and D. Levinson (2010). A Moment of Time: Reliability in Route Choice using Stated Preference. *Journal of Intelligent Transportation Systems 14*(3), 179 –187.

Tipples, J. (2010). Time flies when we read taboo words. *Psychonomic Bulletin & Review 17*, 563–568.

Train, K. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge University Press.

Trueblood, D. L. (1952). Effect of Travel Time and Distance on Freeway Usage. *Highway Research Board* (61), 18–37.

Tseng, Y. and E. Verhoef (2008). Value of time by time of day: a stated-preference study. *Transportation Research Part B 42*, 607618.

Tseng, Y., E. Verhoef, G. de Jong, M. Kouwenhoven, and T. van der Hoorn (2009). A pilot study into the perception of unreliability of travel times using in-depth interviews. *Journal of Choice Modelling 2*(1), 8–28.

Tversky, A. and D. Kahneman (1974). Judgment under Uncertainty: Heuristics and Biases. *Science 185*(4157), 1124–1131.

Tversky, A. and D. Kahneman (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty 5*, 297–323.

Varian, H. (1978). *Microeconomic Analysis* (3rd ed.). W. W. Norton and Company, New York.

Vaziri, M. and T. Lam (1983). Perceived Factors Affecting Driver Route Decisions. *Journal of Transportation Engineering 109*(2), 297311.

Vickrey, W. (1969). Congestion theory and transport investment. *The American Economic Review 59*, 251–260.

Vickrey, W. (1973). Pricing, metering, and efciently using urban transportation facilities. *Highway Research Record 476*, 36–48.

Vierordt, K. (1868). *Der Zeitsinn nach Versuchen*. Tubingen: H Laupp.

Von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Wardman, M. (1998). The value of travel time: a review of British evidence. *Journal of Transport Economics and Policy 32*, 235–316.

Wardman, M. (2001). A review of british evidence on time and service quality valuations. *Transportation Research Part E 37*(2-3), 107–128.

Wardman, M. (2004). Public Transport Values of Time. *Transport Policy 11*, 363–377.

Wong, H. and J. Sussman (1973). Dynamic travel time estimation on highway networks. *Transportation Research 7*, 355–370.

Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press.

Xie, F. and D. Levinson (2007). Measuring the Structure of Road Networks. *Geographical Analysis 39*, 336–356.

Yan, J. (2002). *Heterogeneity in motorists' preferences for travel time and time reliability: empirical finding from multiple survey data sets and its policy implications*. Ph. D. thesis, University of California, Irvine (USA).

Zakay, D. (1992). On prospective time estimation, time relevance and temporal uncertainty. In F. Macar, I. Poutas, and W. Friedman (Eds.), *Time, cognition and action*, pp. 109–119. Kluwer Academic Publishers.

Zamparini, L. and A. Reggiani (2007a). Freight Transport and the Value of Travel Time Savings: A Meta-analysis of Empirical Studies. *Transport Reviews 27*(5), 621–636.

Zamparini, L. and A. Reggiani (2007b). Meta-Analysis and the Value of Travel Time Savings: A Transatlantic Perspective in Passenger Transport. *Networks and Spatial Economics 7*(4), 377–396.

Zhang, L. and D. Levinson (2008). Determinants of Route Choice and the Value of Traveler Information: A Field Experiment. *Transportation Research Record: Journal of the Transportation Research Board 2086*, 81–92.

Zhu, S. (2010). *The Roads Taken: Theory and Evidence on Route Choice in the wake of the I-35W Mississippi River Bridge Collapse and Reconstruction*. Ph. D. thesis, University of Minnesota, Twin Cities (USA).

Appendix A: Filtering survey for subject recruitment

Questions about your background, transportation choices and preferences

Do you currently have a valid Minnesota driver's license?

- No
- Yes

Are you between 25 to 65 years of age?

- No
- Yes

What is your gender?

- Female
- Male

Do you drive to downtown Minneapolis to work at least 4 days a week?

- No
- Yes

If yes, what is your main work location?

Address

City

State

Zip

What is your normal departure time from HOME?

Hours:Minutes

What is your normal departure time from WORK?

Hours:Minutes

Where do you reside?

Address

City

State

Zip

Which mode of transportation do you use most often to get to work?

- Drive alone (Automobile, Light truck, etc.)
- Carpool/Vanpool driver
- Carpool/Vanpool passenger
- Bus /Light Rail /Park and ride
- Motorcycle
- Bicycle
- Walk
- Other, Please specify

From which resources did you hear about this study?

- On-line advertisement at Craigslist
- On-line advertisement at City Pages
- Newspaper advertisement in City Pages
- Flyer at downtown parking ramp
- Flyer at grocery store
- Flyer at county or city libraries
- From friends, co-workers, or family members
- Email

If you drive a vehicle to work, what is the car you most frequently drive?

Are you willing to allow a GPS device to be installed in your vehicle for the duration of the study? (The GPS device will be placed on your vehicle's dashboard. The installation will not alter your vehicle in any way. All data will be kept confidential and anonymous and will be used for the purposes of this study only.)

- No
- Yes

Submit

Page 1 of 2

Contact information

Thank you very much for your interest in this study. Please provide your contact information below. We will contact you if you are eligible to participate in this study. Your name and contact information will not be shared with anyone else and they will not be spammed. Your responses to the screening questions will be destroyed after the selection process.

First name:

Last Name:

Email:

Work phone number:

Home phone number:

Mobile phone number:

Submit

Page 2 of 2

Appendix B: Periodic web survey for subjects

Questions regarding your travel experience

1. Which day of the week are you evaluating for your travel experience?

- Monday
- Tuesday
- Wednesday
- Thursday
- Friday

2. Did you go to work today?

- Yes
- No (If no, please skip to the question 41)

3. Did you cross the Mississippi River on your way to work?

- Yes
- No (If no, please go to question 5)

4. If yes, which bridge did you use?

- I-35W Mississippi River Bridge
- I-94 Mississippi River Bridge
- I-694 Mississippi River Bridge
- Lowry Ave Bridge
- Broadway Ave Bridge
- 8th Ave Bridge
- Hennepin Ave Bridge
- 3rd Ave Bridge
- Cedar Ave Bridge (10th Ave)
- Washington Ave Bridge
- Franklin Ave Bridge
- Lake Street/Marshall Ave. Bridge
- Ford Parkway Bridge
- Other, please specify

5. Did you decide your route to work before leaving home or en route?

- Before leaving home
- En route

6. What was the MOST important reason you chose this route?

- Travel time
- Travel time predictability
- Cost (including tolls)
- Distance
- Avoid stop signs or traffic lights
- Avoid ramp meters
- Follow detour signs
- Realtime GPS guidance
- Convenience for shopping
- Drop off spouse
- Drop off children
- Aesthetics of roads
- Others, please specify

7. What was the SECOND MOST important reason you chose this route?

- Travel time
- Travel time predictability
- Cost (including tolls)
- Distance
- Avoid stop signs or traffic lights
- Avoid ramp meters
- Follow detour signs
- Realtime GPS guidance
- Convenience for shopping
- Drop off spouse
- Drop off children
- Aesthetics of roads
- Others, please specify

8. How much flexibility did you have in the time you had to arrive at work?

- Had to be there at a specified time
- Had to be there within 10 minutes of a specified time
- Within 20 minutes
- Within 30 minutes
- Within 45 minutes
- Within 60 minutes
- Or had more flexibility in the arrival time than that.

9. How long was your trip to work?

☐ minutes

10. What did you EXPECT the actual travel time to work would be?

☐ minutes

11. Did you check traffic conditions from the following resources before leaving for work? If yes, how accurate was the travel information you obtained? (Check all that apply)

☐ TV	Not at all accurate	1	2	3	4	5	6	7	Very accurate
☐ Radio	Not at all accurate	1	2	3	4	5	6	7	Very accurate
☐ Website	Not at all accurate	1	2	3	4	5	6	7	Very accurate
☐ Call 511	Not at all accurate	1	2	3	4	5	6	7	Very accurate
☐ Other	Not at all accurate	1	2	3	4	5	6	7	Very accurate

If you chose other, please specify the information resource you checked?

12. Did you make any stops or side trips on your trip to WORK?

○ Yes
○ No (If no, please go to the question 21)

13. If Yes, which of the following best describes the type of stops you made?

☐ Dropped child off at day care
☐ Dropped someone else off
☐ Picked people up
☐ Took care of personal business, like shopping
☐ Did a work-related activity
☐ Visited friends or other family members
☐ Others, please specify

14. Did you plan for this stop before leaving home?

○ Yes

○ No

15. How often do you make this stop?

○ Every day
○ Several times a week
○ Once a week
○ Once a month
○ Less than once a month

16. Who was with you during this side trip?

☐ Spouse
☐ Child
☐ Colleagues
☐ Other friends

17. Was this stop discretionary (you could make this trip at any time) or mandatory (you had to make it at this time)?

○ Discretionary
○ Mandatory

18. Could anyone in the household have made this side trip for you?

○ Yes
○ No

19. If the price of gas were $5 a gallon, would you still make this side trip?

○ Yes
○ No

20. If the price of gas were $6 a gallon, would you still make this side trip?

○ Yes
○ No

21. Please describe your travel experience on this trip to work?

	○	○	○	○	○	○	○	
Not at all congested	1	2	3	4	5	6	7	Very congested
	○	○	○	○	○	○	○	
Not at all stressful	1	2	3	4	5	6	7	Very stressful

****

Please anwswer some questions regarding your trip to HOME

22. Did you cross the Mississippi River on your way home?

- ○ Yes
- ○ No

23. If yes, which bridge did you use?

- ○ I-35W Mississippi Bridge
- ○ I-94 Mississippi Bridge
- ○ I-694 Mississippi Bridge
- ○ Lowry Ave Bridge
- ○ Broadway Ave Bridge
- ○ 8th Ave Bridge
- ○ Hennepin Ave Bridge
- ○ 3rd Ave Bridge
- ○ Cedar Ave Bridge (10th Ave)
- ○ Washington Ave Bridge
- ○ Franklin Ave Bridge
- ○ Lake Street/Marshall Ave. Bridge
- ○ Ford Parkway Bridge
- ○ Others, please specify

24. Did you decide your route to home before leaving from work or en route?

- ○ Before leaving from work
- ○ En route

25. What was the MOST important reason you chose this route?

- ○ Travel time
- ○ Travel time predictability
- ○ Cost (including tolls)
- ○ Distance
- ○ Avoid stop signs or traffic lights
- ○ Avoid ramp meters
- ○ Follow detour signs
- ○ Realtime GPS guidance
- ○ Convenience for shopping
- ○ Drop off spouse
- ○ Drop off children

- ○ Aesthetics of roads
- ○ Others, please specify [ ]

26. What was the SECOND MOST important reason you chose this route?

- ○ Travel time
- ○ Travel time predictability
- ○ Cost (including tolls)
- ○ Distance
- ○ Avoid stop signs or traffic lights
- ○ Avoid ramp meters
- ○ Follow detour signs
- ○ Realtime GPS guidance
- ○ Convenience for shopping
- ○ Drop off spouse
- ○ Drop off children
- ○ Aesthetics of roads
- ○ Others, please specify [ ]

27. How much flexibility did you have in the time you had to arrive HOME?

- ○ Had to be there at a specified time
- ○ Had to be there within 10 minutes of a specified time
- ○ Within 20 minutes
- ○ Within 30 minutes
- ○ Within 45 minutes
- ○ Within 60 minutes
- ○ Or had more flexibility in the arrival time than that.

28. How long was your trip to home?

[ ] minutes

29. What did you EXPECT the actual travel time to home would be?

[ ] minutes

30. Did you check traffic conditions from the following resources before leaving FROM work? If yes, how accurate was the travel information you obtained? (Check all that apply)

		○	○	○	○	○	○	○	
☐ TV	Not at all accurate	1	2	3	4	5	6	7	Very accurate
		○	○	○	○	○	○	○	

☐ Radio	Not at all accurate	1	2	3	4	5	6	7	Very accurate
		○	○	○	○	○	○	○	
☐ Website	Not at all accurate	1	2	3	4	5	6	7	Very accurate
		○	○	○	○	○	○	○	
☐ Call 511	Not at all accurate	1	2	3	4	5	6	7	Very accurate
		○	○	○	○	○	○	○	
☐ Other	Not at all accurate	1	2	3	4	5	6	7	Very accurate

If you chose other, please specify the information resource you checked?

31. Did you make any stops or side trips on your trip HOME?

○ Yes
○ No (If no, please go to question 40)

32. If Yes, which of the following best describes the type of stops you made?

☐ Picked child up at day care
☐ Picked someone else up
☐ Dropped people off
☐ Took care of personal business, like shopping
☐ Did a work-related activity
☐ Visited friends or other family members
☐ Others, please specify

33. Did you plan for this stop before leaving from work?

○ Yes
○ No

34. How often do you make this stop or side trip?

○ Every day
○ Several times a week
○ Once a week
○ Once a month
○ Less than once a month

35. Who was with you during this side trip?

☐ Spouse

- ☐ Child
- ☐ Colleagues
- ☐ Other friends

36. Was this stop discretionary (you could make this trip at any time) or mandatory (you had to make it at this time)?

- ○ Discretionary
- ○ Mandatory

37. Could anyone in the household have made this side trip for you?

- ○ Yes
- ○ No

38. If the price of gas were $5 a gallon, would you still make this side trip?

- ○ Yes
- ○ No

39. If the price of gas were $6 a gallon, would you still make this side trip?

- ○ Yes
- ○ No

40. Please describe your travel experience on this trip?

	○	○	○	○	○	○	○	
Not at all congested	1	2	3	4	5	6	7	Very congested
	○	○	○	○	○	○	○	
Not at all stressful	1	2	3	4	5	6	7	Very stressful

41. Have you been involved in any crashes in the past week?

- ○ No
- ○ Yes

If Yes, please describe it here

42. Has your vehicle broken down in the past week?

- ○ No
- ○ Yes

If Yes, please describe it here

Submit

Appendix C: Final web survey for subjects

Final survey

1. How would you describe the current condition of the **I-94 Mississippi River Bridge** with regard to the following aspects? (If you are not sure, please choose the option "Not sure" and go directly to the next question)

Congestion level:

	○	○	○	○	○	○	○	
Not at all congested	1	2	3	4	5	6	7	Extremely congested

☐ Not sure

Travel time predictability:

	○	○	○	○	○	○	○	
Not at all predictable	1	2	3	4	5	6	7	Very predictable

☐ Not sure

Ease of driving:

	○	○	○	○	○	○	○	
Very difficult	1	2	3	4	5	6	7	Very easy

☐ Not sure

Pleasantness:

	○	○	○	○	○	○	○	
Very unpleasant	1	2	3	4	5	6	7	Very pleasant

☐ Not sure

2. How does the current condition of the **I-94 Mississippi River Bridge** differ from what it was before the reopening of the I-35W Mississippi River Bridge (two months ago) with regard to the following aspects? (1 = Much better, 4 = No change, 7 = Much worse; if you are not sure, please choose "Not sure.")

Congestion level:

	○	○	○	○	○	○	○	
Much better	1	2	3	4	5	6	7	Much worse

☐ Not sure

Travel time predictability:

	○	○	○	○	○	○	○	
Much better	1	2	3	4	5	6	7	Much worse

☐ Not sure

Ease of driving:

	○	○	○	○	○	○	○	
Much better	1	2	3	4	5	6	7	Much worse

☐ Not sure

Pleasantness:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

3. How would you describe the current condition of the **I-694 Mississippi River Bridge** with regard to the following aspects? (If you are not sure, please choose the option "Not sure" and directly go to the next question)

Congestion level:

Not at all congested 1 2 3 4 5 6 7 Extremely congested

☐ Not sure

Travel time predictability:

Not at all predictable 1 2 3 4 5 6 7 Very predictable

☐ Not sure

Ease of driving:

Very difficult 1 2 3 4 5 6 7 Very easy

☐ Not sure

Pleasantness:

Very unpleasant 1 2 3 4 5 6 7 Very pleasant

☐ Not sure

4. How does the current condition of the **I-694 Mississippi River Bridge** differ from what it was before the reopening of the I-35W Mississippi River Bridge (two months ago) with regard to the following aspects? (1 = Much better, 4 = No change, 7 = Much worse; if you are not sure, please choose "Not sure.")

Congestion level:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

Travel time predictability:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

Ease of driving:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Pleasantness:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

5. How would you describe the current condition of the **Hennepin Avenue Bridge** crossing the Mississippi River with regard to the following aspects? (If you are not sure, please choose the option "Not sure" and directly go to the next question)

Congestion level:

Not at all congested 1 2 3 4 5 6 7 Extremely congested

Not sure

Travel time predictability:

Not at all predictable 1 2 3 4 5 6 7 Very predictable

Not sure

Ease of driving:

Very difficult 1 2 3 4 5 6 7 Very easy

Not sure

Pleasantness:

Very unpleasant 1 2 3 4 5 6 7 Very pleasant

Not sure

6. How does the current condition of the **Hennepin Avenue Bridge** differ from what it was before the reopening of the I-35W Mississippi River Bridge (two months ago) with regard to the following aspects? (1 = Much better, 4 = No change, 7 = Much worse; if you are not sure, please choose "Not sure.")

Congestion level:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Travel time predictability:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Ease of driving:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Pleasantness:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

7. How would you describe the current condition of the **3rd Avenue Bridge** crossing the Mississippi River with regard to the following aspects? (If you are not sure, please choose the option "Not sure" and directly go to the next question)

Congestion level:

Not at all congested 1 2 3 4 5 6 7 Extremely congested

Not sure

Travel time predictability:

Not at all predictable 1 2 3 4 5 6 7 Very predictable

Not sure

Ease of driving:

Very difficult 1 2 3 4 5 6 7 Very easy

Not sure

Pleasantness:

Very unpleasant 1 2 3 4 5 6 7 Very pleasant

Not sure

8. How does the current condition of the **3rd Avenue Bridge** differ from what it was before the reopening of the I-35W Mississippi River Bridge (two months ago) with regard to the following aspects? (1 = Much better, 4 = No change, 7 = Much worse; if you are not sure, please choose "Not sure.")

Congestion level:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Travel time predictability:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

Ease of driving:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

Pleasantness:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

9. How would you describe the current condition of the **Cedar Avenue Bridge (10th Avenue)** crossing the Mississippi River with regard to the following aspects? (If you are not sure, please choose the option "Not sure" and directly go to the next question)

Congestion level:

Not at all congested 1 2 3 4 5 6 7 Extremely congested

☐ Not sure

Travel time predictability:

Not at all predictable 1 2 3 4 5 6 7 Very predictable

☐ Not sure

Ease of driving:

Very difficult 1 2 3 4 5 6 7 Very easy

☐ Not sure

Pleasantness:

Very unpleasant 1 2 3 4 5 6 7 Very pleasant

☐ Not sure

10. How does the current condition of the **Cedar Avenue Bridge (10th Avenue)** differ from what it was before the reopening of the I-35W Mississippi River Bridge (two months ago) with regard to the following aspects? (1 = Much better, 4 = No change, 7 = Much worse; if you are not sure, please choose "Not sure.")

Congestion level:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

Travel time predictability:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Ease of driving:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Pleasantness:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

11. How would you describe the current condition of the **Washington Avenue Bridge** crossing the Mississippi River with regard to the following aspects? (If you are not sure, please choose the option "Not sure" and directly go to the next question)

Congestion level:

Not at all congested 1 2 3 4 5 6 7 Extremely congested

Not sure

Travel time predictability:

Not at all predictable 1 2 3 4 5 6 7 Very predictable

Not sure

Ease of driving:

Very difficult 1 2 3 4 5 6 7 Very easy

Not sure

Pleasantness:

Very unpleasant 1 2 3 4 5 6 7 Very pleasant

Not sure

12. How does the current condition of the **Washington Avenue Bridge** differ from what it was before the reopening of the I-35W Mississippi River Bridge (two months ago) with regard to the following aspects? (1 = Much better, 4 = No change, 7 = Much worse; if you are not sure, please choose "Not sure.")

Congestion level:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Travel time predictability:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Ease of driving:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Pleasantness:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

13. How would you describe the current condition of **Highway 280** with regard to the following aspects? (If you are not sure, please choose the option "Not sure" and directly go to the next question)

Congestion level:

Not at all congested 1 2 3 4 5 6 7 Extremely congested

Not sure

Travel time predictability:

Not at all predictable 1 2 3 4 5 6 7 Very predictable

Not sure

Ease of driving:

Very difficult 1 2 3 4 5 6 7 Very easy

Not sure

Pleasantness:

Very unpleasant 1 2 3 4 5 6 7 Very pleasant

Not sure

14. How does the current condition of **Highway 280** differ from what it was before the reopening of the I-35W Mississippi River Bridge (two months ago) with regard to the following aspects? (1 = Much better, 4 = No change, 7 = Much worse; if you are not sure, please choose "Not sure.")

Congestion level:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

Travel time predictability:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

Ease of driving:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

Pleasantness:

Much better 1 2 3 4 5 6 7 Much worse

☐ Not sure

15. How would you describe the current condition of the **I-35W Mississippi River Bridge** with regard to the following aspects? (If you are not sure, please choose the option "Not sure" and directly go to the next question)

Congestion level:

Not at all congested 1 2 3 4 5 6 7 Extremely congested

☐ Not sure

Travel time predictability:

Not at all predictable 1 2 3 4 5 6 7 Very predictable

☐ Not sure

Ease of driving:

Very difficult 1 2 3 4 5 6 7 Very easy

☐ Not sure

Pleasantness:

Very unpleasant 1 2 3 4 5 6 7 Very pleasant

☐ Not sure

16. How does the current condition of the **I-35W Mississippi River Bridge** differ from what it was **before its collapse** one year ago with regard to the following aspects? (1 = Much better, 4 = No change, 7 = Much worse; if you are not sure, please choose "Not sure.")

Congestion level:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Travel time predictability:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Ease of driving:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Pleasantness:

Much better 1 2 3 4 5 6 7 Much worse

Not sure

Next

Final survey

1. Did you change your usual routes from home to work after the reopening of the I-35W Bridge?

- Yes
- No

If you responded Yes, continue to the next question. If you responded No, please skip the next question.

2. What was the **most** important reason you changed your route after the I-35W Bridge reopened?

- The route I followed before the reopening of I-35W Bridge is more congested now.
- The new route has a shorter travel distance.
- The new route has a shorter travel time.
- The travel time of the new route is more reliable (predictable).
- Others, please specify

3. Did you try alternative routes other than your usual routes after the I-35W Bridge reopened?

- Yes
- No

If you responded No, continue to the next question. If you responded Yes, please skip the next question.

4. What was the most important reason for you to stick to your usual routes without trying alternatives?

- There is no real alternative for my route to work.
- I do not know if there are alternative routes and do not want to bother.
- The alternative routes are not likely to be better off.
- The time and effort of trying alternatives outweighs possible time savings.
- Others, please specify

5. Please rank your route preferences for driving to WORK.

Most preferred	1	2	3	Least preferred
I-35W Mississippi Bridge	○	○	○	
I-94 Mississippi Bridge	○	○	○	
I-694 Mississippi Bridge	○	○	○	

Hennepin Avenue Bridge	○	○	○
3rd Avenue Bridge	○	○	○
Cedar Avenue Bridge (10th Avenue)	○	○	○
Washington Avenue Bridge	○	○	○
Franklin Avenue Bridge	○	○	○
Others	○	○	○

If you chose others, please specify

6. Please rank the importance of the following factors (top three) when you choose a route to WORK.

Most important	1	2	3	Least important
Travel time	○	○	○	
Distance	○	○	○	
Travel time predictability	○	○	○	
Cost (including tolls)	○	○	○	
Convenience for shopping	○	○	○	
Drop off spouse	○	○	○	
Drop off children	○	○	○	
Aesthetics of route	○	○	○	
Others	○	○	○	

If you chose others, please specify

7. Please rank your route preferences for driving HOME.

Most preferred	1	2	3	Least preferred
I-35W Mississippi Bridge	○	○	○	
I-94 Mississippi Bridge	○	○	○	
I-694 Mississippi Bridge	○	○	○	
Hennepin Avenue Bridge	○	○	○	
3rd Avenue Bridge	○	○	○	
Cedar Avnuee Bridge (10th Avenue)	○	○	○	
Washington Avenue Bridge	○	○	○	

	1	2	3
Franklin Avenue Bridge	○	○	○
Others	○	○	○

If you chose others, please specify

8. Please rank the importance of the following factors (top three) when you choose a route HOME.

Most important	1	2	3	Least important
Travel time	○	○	○	
Distance	○	○	○	
Travel time predictability	○	○	○	
Cost (including tolls)	○	○	○	
Convenience for shopping	○	○	○	
Drop off spouse	○	○	○	
Drop off children	○	○	○	
Aesthetics of route	○	○	○	
Others	○	○	○	

If you chose others, please specify

9. What other activities do you engage in which require you to make trips that cross the Mississippi River? (Choose all that apply.)

- ☐ Childcare
- ☐ Quick stop
- ☐ Shopping
- ☐ Visit friends/Relatives
- ☐ Personal business
- ☐ Eat meal outside of home
- ☐ Entertainment/Recreational/Fitness
- ☐ Civic/Religious
- ☐ Pick up/Drop off
- ☐ With another person at their activity
- ☐ Others, please specify

10. Which of those activities affects your route choice the most?

- ○ Childcare
- ○ Quick stop
- ○ Shopping

- Visit friends/Relatives
- Personal business
- Eat meal outside of home
- Entertainment/Recreational/Fitness
- Civic/Religious
- Pick up/Drop off
- With another person at their activity
- Others, please specify

11. Please rank your route preferences for that purpose.

Most preferred	1	2	3	Least preferred
I-35W Mississippi Bridge	○	○	○	
I-94 Mississippi Bridge	○	○	○	
I-694 Mississippi Bridge	○	○	○	
Hennepin Avenue Bridge	○	○	○	
3rd Avenue Bridge	○	○	○	
Cedar Avenue Bridge (10th Ave)	○	○	○	
Washington Avenue Bridge	○	○	○	
Franklin Avenue Bridge	○	○	○	
Others	○	○	○	

If you chose others, please specify

12. Please rank the importance of the following factors (top three) when you choose a route for that purpose?

Most important	1	2	3	Least important
Travel time	○	○	○	
Distance	○	○	○	
Travel time predictability	○	○	○	
Cost (including tolls)	○	○	○	
Convenience for shopping	○	○	○	
Drop off spouse	○	○	○	
Drop off children	○	○	○	
Aesthetics of route	○	○	○	

Others

If you chose others, please specify

The following questions are about your travel preferences after the I-35W Bridge collapse

13. Did you change your usual routes from home to work after the I-35W Bridge Collapse one year ago?

- Yes
- No

If you responded Yes, continue to the next question. If you responded No, please skip the next question.

14. What was the **most** important reason you changed your route after the I-35W Bridge Collapse?

- Routes or ramp closed because of the bridge collapse.
- The traffic condition on the usual route before the bridge collapse was much worse.
- The traffic condition on new route was better than the usual route before the bridge collapse.
- The travel time of the new route was more reliable (predictable).
- Others, please specify

15. Did you try alternative routes other than your usual route after the bridge collapse?

- Yes
- No

If you responded No, continue to the next question. If you responded Yes, please skip the next question.

16. What is the **most** important reason for you to stick to your usual route without trying alternatives after the bridge collaspe?

- There is no real alternative for my route to work.
- I do not know if there are alternative routes and do not want to bother.
- The alternative routes are not likely to be better off.
- The time and efforts for trying alternatives outweight possible time savings.
- Others, please specify

17. Did you make fewer crossing-river trips after the bridge collapse?

- Yes
- No

If you responded Yes, continue to the next question. If you responded No, please skip the next question.

18. If yes, how many trips did you cancel or consolidate with other trips?

- Several trips per day
- Several trips a week
- Once a week
- Once a month
- Less than once a month

19. Did you change your departure time from home to work after the bridge collapse?

- Yes
- No

If Yes, by how much?

minutes

- earlier
- later

20. Could you please comment on the impacts of the I-35W Bridge collapse regarding your travel pattern?

Next

Page 1b of 14

Final Questions

For the following questions please choose a number from 1 – 7 that represents your response. For example, an answer of 1 means that you never worry and an answer of 7 means that you always worry.

1. Do you sometimes worry about driving on bridges or overpasses?

- Yes
- No

If yes, please answer the following question. If no, continue to Question 2.
How often do you worry?

Never	1	2	3	4	5	6	7	Always

2. Do you sometimes worry about driving under a bridge or overpass?

- Yes
- No

If yes, please answer the following question. If no, continue to Question 3.
How often do you worry?

Never	1	2	3	4	5	6	7	Always

3. Do you sometimes worry that a bridge or overpass might collapse when you are driving on it?

- Yes
- No

If yes, please answer the following question. If no, continue to Question 4.
How often do you worry?

Never	1	2	3	4	5	6	7	Always

4. Do you sometimes worry that a bridge or overpass might collapse when you are driving under it?

- Yes
- No

If yes, please answer the following question. If no, continue to Question 5.
How often do you worry?

	○	○	○	○	○	○	○	
Never	1	2	3	4	5	6	7	Always

5. Before the I-35W Bridge collapsed, did you sometimes worry that a bridge or overpass might collapse while you were driving on it?

○ Yes
○ No

If yes, please answer the following question. If no, continue to Question 6.
How often do you worry?

	○	○	○	○	○	○	○	
Never	1	2	3	4	5	6	7	Always

6. Before the I-35W Bridge collapsed, did you sometimes worry that a bridge or overpass might collapse while you were driving under it?

○ Yes
○ No

If yes, please answer the following question. If no, continue to Question 7.
How often do you worry?

	○	○	○	○	○	○	○	
Never	1	2	3	4	5	6	7	Always

7. If you worry about driving on bridges and overpasses, or under them, does this affect how you drive or where you drive?

○ Yes
○ No

If yes, please comment below:

8. What is the highest grade or year of school that you have completed?

- 11th grade or less
- High school graduate
- Associate degree
- Bachelors degree
- Masters degree
- Doctoral degree

9. What is your age?

- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65+

10. What is the total annual income for your household, when you consider the income of all employed individuals?

- $30,000 or less
- $30,000 to $49,999
- $50,000 to $74,999
- $75,000 to $99,999
- $100,000 to $124,999
- $125,000 to $149,999
- $150,000 or above

11. Which of the following categories best describes your race or ethnic background?

- White or Caucasian
- Black/African American
- Native American
- Hispanic
- Asian
- Mixed race

○ Others

12. How long have you worked at your current work location?

[ ] Years
[ ] Months

13. How long have you lived in your current house/apartment?

[ ] Years
[ ] Months

14. Where would you like the check and gas card you will receive for participating in this study to be mailed?

Payee [ ]
Address [ ]
City [ ]
State [ ]
Zip [ ]

Next

Page 14 of 14